Emotions are thought to be a key factor that critically influences human decision-making. Emotion regulation can help to mitigate emotion related decision biases and eventually lead to a better decision performance. Serious games emerged as a new angle introducing technological methods to learning emotion regulation, where meaningful biofeedback information displays player’s emotional state.

This thesis investigates emotions and the effect of emotion regulation on decision performance. Furthermore, it explores design and evaluation methods for creating serious games where emotion regulation can be learned and practiced.

The scope of this thesis was limited to serious games for emotion regulation training using psychophysiological methods to communicate user’s affective information. Using the psychophysiological methods, emotions and their underlying neural mechanism have been explored. Through design and evaluation of serious games using those methods, effects of emotion regulation have been investigated where decision performance has been measured and analyzed. The proposed metrics for designing and evaluating such affective serious games have been exhaustively evaluated. The research methods used in this thesis were based on both quantitative and qualitative aspects, with true experiment and evaluation research, respectively.

Serious games approach to emotion regulation was investigated. The results suggested that two different emotion regulation strategies, suppression and cognitive reappraisal, are optimal for different decision tasks contexts. With careful design methods, valid serious games for training those different strategies could be produced. Moreover, using psychophysiological methods, underlying emotion neural mechanism could be mapped to provide optimal level of arousal for a certain task.

The results suggest that it is possible to design and develop serious game applications that provide helpful learning environment where decision makers could practice emotion regulation and subsequently improve their decision making.
Design and Evaluation of Affective Serious Games for Emotion Regulation Training

Petar Jerčić
Design and Evaluation of Affective Serious Games for Emotion Regulation Training

Petar Jerčić

Licentiate Dissertation in Game Development

School of Computing
Blekinge Institute of Technology
SWEDEN
Again among the sciences we consider that that science which is desirable in itself and for the sake of knowledge is more nearly Wisdom than that which is desirable for its results.

Abstract

Emotions are thought to be a key factor that critically influences human decision-making. Emotion regulation can help to mitigate emotion related decision biases and eventually lead to a better decision performance. Serious games emerged as a new angle introducing technological methods to learning emotion regulation, where meaningful biofeedback information displays player’s emotional state. This thesis investigates emotions and the effect of emotion regulation on decision performance. Furthermore, it explores design and evaluation methods for creating serious games where emotion regulation can be learned and practiced. The scope of this thesis was limited to serious games for emotion regulation training using psychophysiological methods to communicate user’s affective information. Using the psychophysiological methods, emotions and their underlying neural mechanism have been explored. Through design and evaluation of serious games using those methods, effects of emotion regulation have been investigated where decision performance has been measured and analyzed. The proposed metrics for designing and evaluating such affective serious games have been exhaustively evaluated. The research methods used in this thesis were based on both quantitative and qualitative aspects, with true experiment and evaluation research, respectively. Serious games approach to emotion regulation was investigated. The results suggested that two different emotion regulation strategies, suppression and cognitive reappraisal, are optimal for different decision tasks contexts. With careful design methods, valid serious games for training those different strategies could be produced. Moreover, using psychophysiological methods, underlying emotion neural mechanism could be mapped to provide optimal level of arousal for a certain task. The results suggest that it is possible to design and develop serious game applications that provide helpful learning environment where decision makers could practice emotion regulation and subsequently improve their decision making.
Contents

1 Introduction 1

2 Background 3
   2.1 Related Work .......................... 13
      2.1.1 Serious Games in the Field of Healthcare and Well-being  14
      2.1.2 Serious Games using Biofeedback .......................... 14
      2.1.3 Serious Games in the Context of Financial Decision Making 15
      2.1.4 Summary ........................................... 16

3 Approach 17
   3.1 Aim and Scope ............................ 17
   3.2 RQs ...................................... 18
   3.3 Methodology .............................. 19
   3.4 Threats .................................. 20
   3.5 Contributions ............................ 21
   3.6 Discussion .............................. 26
   3.7 Conclusion ............................... 27
   3.8 Future Work .............................. 29

4 The Future of Brain-Computer Interface for Games and Interaction Design 31
   4.1 Introduction ............................... 32
   4.2 EPOC for Game Design ..................... 32
   4.3 EPOC for Interaction Design ............... 33

5 Evaluating Games Designed to Improve Financial Capability 35
   5.1 Introduction ............................... 36
   5.2 The design and evaluation framework .......... 36
   5.3 A design and evaluation toolkit for financial capability games ... 37
5.3.1 Financial capability De tool ........................................ 38
5.3.2 Behaviour change De tool ........................................ 38
5.3.3 Game design De tool ........................................ 39
5.3.4 Learning with technology De tool ............................ 40
5.4 Conclusions .............................................................. 41

6 A Serious Game Using Physiological Interfaces for Emotion Regulation Training in the Context of Financial Decision-Making 43
6.1 Introduction ............................................................. 45
6.2 Emotions, Emotion Regulation and Decision-Making .......... 47
6.3 Psychophysiological Measurement of Emotions .............. 50
6.4 Game description ....................................................... 51
   6.4.1 Underlying principle ........................................... 51
   6.4.2 Game concept and gameplay ................................ 51
   6.4.3 Game Mechanics ................................................ 54
6.5 Evaluation ............................................................... 57
   6.5.1 Functionality and Heuristic Evaluation .................... 57
   6.5.2 Playtesting Evaluation Method .............................. 58
   6.5.3 Playtesting Evaluation Results .............................. 59
6.6 Discussion and Conclusion ......................................... 60

7 A study of the effect of emotion regulation in a shooting game using bio-feedback from electroencephalography for training in emotion regulation 63
7.1 Introduction ............................................................. 65
7.2 Brain-Computer Interaction ........................................ 65
   7.2.1 Effects of Emotion Regulation ............................... 66
7.3 Method ................................................................. 69
   7.3.1 EEG ............................................................ 69
   7.3.2 The Game ........................................................ 69
   7.3.3 Participants ....................................................... 70
   7.3.4 Questionnaire ................................................... 70
   7.3.5 Procedure ........................................................ 73
7.4 Results ................................................................. 73
7.5 Discussion .............................................................. 75
7.6 Conclusion .............................................................. 79

8 Integrating Biosignals into Information Systems: A NeuroIs Tool for Improving Emotion Regulation 81

9 Identifying the Moment when Demanding Cognitive Load Over- shadows the Arousal Effect on Pupil Dilation During a Stressful
## CONTENTS

<table>
<thead>
<tr>
<th>Task</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>9.2</td>
<td>Material and methods</td>
</tr>
<tr>
<td>9.2.1</td>
<td>Participants</td>
</tr>
<tr>
<td>9.2.2</td>
<td>Materials</td>
</tr>
<tr>
<td>9.2.3</td>
<td>Apparatus</td>
</tr>
<tr>
<td>9.2.4</td>
<td>Procedure</td>
</tr>
<tr>
<td>9.2.5</td>
<td>Data analysis</td>
</tr>
<tr>
<td>9.3</td>
<td>Results</td>
</tr>
<tr>
<td>9.3.1</td>
<td>Baseline values for pupil dilation</td>
</tr>
<tr>
<td>9.4</td>
<td>Parasympathetic vs. sympathetic pathway activation behavior</td>
</tr>
<tr>
<td>9.5</td>
<td>Discussion</td>
</tr>
<tr>
<td>9.6</td>
<td>Conclusions</td>
</tr>
<tr>
<td>9.7</td>
<td>Future Work</td>
</tr>
<tr>
<td>9.8</td>
<td>Acknowledgements</td>
</tr>
</tbody>
</table>

References 105
## List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANS</td>
<td>Autonomous Nervous System</td>
</tr>
<tr>
<td>BCI</td>
<td>Brain-Computer Interface</td>
</tr>
<tr>
<td>CNS</td>
<td>Central Nervous System</td>
</tr>
<tr>
<td>DE</td>
<td>The Design and Evaluation Framework</td>
</tr>
<tr>
<td>EEG</td>
<td>Electroencephalography</td>
</tr>
<tr>
<td>EMG</td>
<td>Electromyography</td>
</tr>
<tr>
<td>EP</td>
<td>Evoked Potentials</td>
</tr>
<tr>
<td>ERP</td>
<td>Event-Related Potentials</td>
</tr>
<tr>
<td>ERS/ERD</td>
<td>Event-Related De/Synchronization</td>
</tr>
<tr>
<td>fEMG</td>
<td>Facial Electromyography</td>
</tr>
<tr>
<td>GBL</td>
<td>Game-based Learning</td>
</tr>
<tr>
<td>GSR</td>
<td>Galvanic Skin Response</td>
</tr>
<tr>
<td>HCI</td>
<td>Human-Computer Interaction</td>
</tr>
<tr>
<td>HR</td>
<td>Heart Rate</td>
</tr>
<tr>
<td>HRV</td>
<td>Heart Rate Variability</td>
</tr>
<tr>
<td>IS</td>
<td>Information Systems</td>
</tr>
<tr>
<td>PNS</td>
<td>Parasympathetic Nervous System</td>
</tr>
<tr>
<td>PD</td>
<td>Pupil Dilation</td>
</tr>
<tr>
<td>SNS</td>
<td>Sympathetic Nervous System</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

The mysterious concept of emotion still puzzles scientists. James (1884) presented the still unanswered question: What is an emotion? After all this time the answer has not emerged or converged, with little agreement on where emotion stops and its causes and consequences begin. According to Lang, emotions are dispositions, or states of readiness, that help people and organisms to interact with the environment (Lang, 1995). Such, emotions facilitate our interactions with the socioeconomic environment; they foster meaningful interpersonal interactions, prepare behavioral responses, and enable us to take advantageous decisions (Gross, 2008). There is no clear answer to what is and what is not an emotion. Cole et al. (2004, p. 330) captured the nature of emotions in a statement: ”Emotions are powerful, elusive, dynamic processes. They have the capacity to regulate other processes and to be regulated”. Unfortunately, however, emotions are not always accurately processed. Goldstein (1996) discusses how emotions can get ”out of control”, where in the heat of the moment, human decision makers can be overwhelmed by their emotions and ”under the influence of high levels of arousal” lose control over their actions.

It is vital for decision makers and for the organizations they represent, to adequately process emotions, because emotions particularly affect important decisions. Financial decisions fall into this category (Lucey and Dowling, 2005). There are, however, strong interpersonal differences in the capabilities of adequate emotional processing (Lo and Repin, 2002; Lo et al., 2005). It has been shown that traders and investors with high emotion regulation capabilities perform better in trading (Fenton-O’Creevy et al., 2012b,0). In contrast, decision makers with low emotion regulation capabilities frequently are more prone to
CHAPTER 1. INTRODUCTION

taking disadvantageous decisions with undesirable consequences. Emotion regulation can help to mitigate emotion related decision biases and eventually lead to better decision performance. However, how can a decision maker actually learn to improve his or her emotion regulation capabilities? Unfortunately, traditional learning approaches seem to fail when it comes to improving emotion regulation capabilities (Fenton-O’Creevy et al., 2012a).

**Serious games** emerged as a remedy to a traditional learning approach, providing a new angle introducing technological methods to learning. Serious games are (digital) games used for purposes other than mere entertainment (Susi and Johannesson, 2007). Corti (2006) points out an obvious advantage of serious games in allowing learners to experience situations that are impossible in the real world for reasons of safety, cost, time or logistics. Serious games can have positive impact on the development of a number of different skills (Ellis et al., 2006; Mitchell and Savill-Smith, 2004; Corti, 2006; Squire and Jenkins, 2003; Van Eck, 2006; Rieber, 1996). Having those different skills as defined learned outcomes, one can clearly see why serious games are considered as **Game-based Learning (Gbl)** applications (Corti, 2006).

The interdisciplinary process of creating such a technology-enhanced serious game tools requires an iterative design and development approach. One that is flexible enough to support the complex design and evaluation processes, linking goals acquired from empirical studies and evaluating the tool’s potential to meet those goals (Jonassen et al., 2002). Such need gave rise to a Design and Evaluation (DE) framework which proposes practices, methods and metrics to evaluate serious games (set in the financial context), its design process and meeting design objectives, detailed further in Chapter 5.

This thesis explores game and sensor technologies that could be used to support emotion regulation training and therefore improve decision-making, focusing on the financial context. This approach investigates the role and underlying mechanism of emotions in the decision-making process. These research insights supported development of two serious game approaches for training decision-making in the context of finance. A need to create those approaches in an iterative design and development fashion, gave rise to the DE framework, with its practices, methods and metrics to evaluate serious games set in the financial context. This flexible framework supports their design and evaluation process.

Chapter 2 presents the background and terminology. In Chapter 3, Section 3.2 lists research questions. Section 3.5 concerns the contribution of publications, as well as the author’s involvement. The results are discussed and concluded in Section 3.6. Section 3.8 presents points to future work. Finally, the publications are presented in Chapter 4-9.
Chapter 2

Background

Lang (1995) states that emotions are dispositions, or states of readiness, that help people and organisms to interact with the environment. Russell (1980) generally classified emotions by their independent components, *arousal* and *valence*. In his model, arousal represents the level of excitement whereas valence defines whether the current emotional state is positive or negative, as illustrated in Figure 1. Thus, emotion and their measurements contain a combination of valence and arousal. The concept of *physiological arousal* has been generally validated in models of emotion (Cannon, 1994; Porges, 1995; Morrison, 2001), with one example being the pupil dilation, influenced by physiological arousal, together with attention and interest (Andreassi, 2007).

*Psychophysiology* is concerned with the physiological bases of psychological processes, as a perspective of studying the interface of subjective states and physiological processes. Psychophysiological measures are often used to study emotion, emotional states (Cacioppo and Tassinary, 1990; Ekman et al., 1983; Vyzas and Picard, 1998) and attentional responses to stimuli, during exertion, and increasingly, to better understand cognitive processes.

To measure an emotion, that is a combination of valence and arousal, different *psychophysiological methods* have been used. Psychophysiological methods draw theoretical background from Cognitive Science (Neuroscience, Neuroengineering) and Affective Computing. Affective Computing is a branch of science which studies psychophysiology and its application in digital games, autonomous robotic systems or software applications. Analysis of user’s psychophysiological biosignals creates an interpretation model of the user’s affect, such model could be communicated inside of a digital game, or any other autonomous robotic sys-
Figure 2.1: Emotions in the valence-arousal space
tems or software application. Such games or systems could be aware of player’s emotional states and can provide meaningful biofeedback information.

Many modality measures are a part of modern psychophysiology (Figure 2) from which the most interesting for my research are brain waves (electroencephalography, EEG); further supported by measures of skin conductance (galvanic skin response, GSR), cardiovascular measures (heart rate, HR; heart rate variability, HRV), muscle activity (electromyography, EMG), and changes in pupil diameter in regards to cognition and emotion (pupillometry or pupil dilation, PD). These measures may provide key information regarding the intensity and quality of an individual’s internal experience and affect; furthermore, they can easily be digitized and may eventually be unobtrusively (noninvasively) monitored, making them very accessible to pattern recognition techniques (Scheirer et al., 2002).

Electroencephalography (EEG) is the recording of electrical activity along the scalp. Methods like evoked potentials (EPs) and Event-Related Potentials (ERPs) average EEG responses which are time-locked to the presentation of a stimulus of some sort (visual, somatosensory, or auditory). They offer us the possibility of analyzing EEG responses. This thesis is mostly concerned with Event Related De/Synchronization (ERS/ERD) where power of certain frequencies gives information on the affect person is feeling. Regardless of the analysis goal, most of the signal processing methods (Bashashati et al., 2007) and classification algorithms (Lotte and Congedo, 2007) for EEG have been developed in the context of building Brain Computer Interfaces (BCIs), and researchers are constantly seeking ways for developing new approaches for recognizing affective states from EEG and other physiological signals. These techniques are used in cognitive science, cognitive psychology, and psychophysiological research. A Brain Computer Interface (BCI) is a direct communication pathway between the brain and an external device, a system which allows someone to communicate information about their mental state without the use of the peripheral nervous system (Calvo and D’Mello, 2010).

Galvanic Skin Response (GSR) is a method of measuring the electrical conductance of the skin (Dawson et al., 2011), where the skin momentarily becomes a better conductor of electricity when either external or internal physiologically arousing stimuli occur and excite the eccrine sweat glands under the skin, controlled by the sympathetic nervous system (SNS). Since there is a relationship between sympathetic activity and emotional arousal, GSR has been closely linked to both emotion and attention (Scheirer et al., 2002), although it is used as an indication of psychological or physiological arousal, one cannot identify which specific emotion is being elicited.
CHAPTER 2. BACKGROUND

Figure 2.2: Example of four physiological signals measured from an actress while she intentionally expressed anger (left) and grief (right). From top to bottom: Electromyogram (microvolts), blood volume pressure (percent reflectance), galvanic skin conductivity (microSiemens), and respiration (percent maximum expansion). The signals were sampled at 20 samples a second. Each box shows 100 seconds of response. The segment shown here are visibly different for the two emotions, which was not true in general. (Vyzas and Picard, 1998)
**Heart rate** (HR) represents the number of heartbeats per unit of time. Under the influence of SNS and the *parasympathetic nervous system* (PNS), the heart rate, if measured precisely enough, can be used to extract **heart-rate variability** (HRV) which represents variability between heart beats intervals. In the field of psychophysiology, there is great interest in HRV since it is related to emotional arousal/valence (Rowe et al., 1998; Scheirer et al., 2002).

Pupillary response, the variation of the pupil size in response to a stimulus under the influence of *autonomous nervous system* (ANS). It is a physiological response that varies the size of the pupil, either resulting in constriction (SNS) or dilation (PNS), via activation or deactivation of the iris dilator muscles (Ellis, 1981). The pupil dilates in response to extreme emotional situations such as fear, or to contact of a sensory nerve, such as pain. Previous studies have found that pupil size variation is also related to both cognitive and affective cues, fear and pain being an example of pupil dilation response stimuli. If emotions and pupil size variation were reliably associated with each other, then current eye-tracking technology would offer a possibility for unobtrusive monitoring of emotion-related reactions because no sensors would need to be attached to the user. In order to be able to evaluate the possibilities of using pupil size measurement for detecting the users’ emotional responses, we need to understand how emotions and pupil size variation relate to each other (Partala and Surakka, 2003).

**Electromyography** (EMG) is a technique for evaluating and recording the electrical activity produced by skeletal muscles. *Facial electromyography* (fEMG) refers to measurements of the muscles of the face indicating an emotional reactions (Dimberg, 1990). Studies have found that activity of the corrugator muscle (frowns) and the zygomatic muscle (smiling) is said to be associated with emotional state. Facial EMG has been used as a technique to distinguish and track positive and negative emotional reactions to a stimulus as they occur (Wolf et al., 2005).

More recently, psychophysicists have been interested in the central nervous system (CNS), exploring cortical brain potentials such as the many types of ERPs, brain waves, and utilizing advanced technology mentioned in the previous paragraph. Cognitive Psychology and Psychophysiology are closely related to the field of Neuroscience, which primarily concerns itself with relationships between psychological events and brain responses.

The extent interplay between SNS or PNS activities provides a measure of a subject’s emotion regulation capabilities. Those two synergistic pathways control pupil size and dynamics, operating on the smooth muscles of the pupil. The PNS pathway is mediated with the efferent pathway originating in the Edinger-Westphal oculomotor complex in the midbrain and innervates the sphincter which
is the circular muscle responsible for constriction, as seen in the reflex reaction to light. Inhibition of Edinger-Westphal complex results in relaxation of the sphincter muscles and, thus, significant dilation. The SNS pathway, mediated by the hypothalamus, innervates the radial dilator muscle of the iris responsible for dilation (Steinhauer et al., 2004; Privitera et al., 2008). Similarly, Berntson et al. (2007) states that the heart rate is also a measure of activity of both the SNS and PNS part of ANS. Correspondingly, SNS activity tends to increase heart rate, while PNS activity decreases it. The described mechanism has been illustrated on the schematic view on Figure 3.

Figure 2.3: Model of the ANS control pathways for the pupil and the heart

*Cognitive science* is an interdisciplinary scientific study of the mind and its processes regarding cognition, perception, language, memory, reasoning and emotion (Friedenberg and Silverman, 2005). Furthermore, it also explores on how is information represented and transformed in human/animal behavior, nervous systems or machines. As the field is highly interdisciplinary, research often cuts across multiple areas of study, drawing on research methods previously men-
tioned, such as psychology, neuroscience, computer science and systems theory. By measuring behavioral responses (by means of EEG and pupillary responses) to different stimuli, one can understand something about a variety of cognitive processes and how those stimuli are processed.

While Neuroscience is the scientific study of the nervous system, Cognitive neuroscience is an academic field concerned with the scientific study of biological substrates underlying cognition, with a specific focus on the neural substrates of mental processes (Squire et al., 2008). It explores the underlying neural mechanism of psychological/cognitive functions in the brain. At the cognitive level, cognitive neuroscience addresses questions of how psychological functions are produced by neural circuitry. Neural imaging and sophisticated experimental techniques from cognitive psychology, allows addressing abstract questions such as how human cognition and emotion are mapped to specific neural substrates.

Brain-Computer Interface (BCI) methodology uses EEG to extract signals providing activation of different brain regions and frequency of the brain rhythms. Brain-Computer Interface is the neuroscientific interaction technology in the field of Human-Computer Interaction (HCI) (van Erp et al., 2011). BCI, as defined by Schalk et al. (2008), has four criteria: (i) the recording device must rely on signals recorded from the brain; (ii) there must be at least one recordable brain signal that the user can learn how to intentionally manipulate; (iii) real-time processing must be available; and (iv) the user must obtain feedback. Note that this does not exclude systems that utilize additional input (e.g. mouse or keyboard) as well as recorded data from the brain. BCI methods hold a noticeably reduced response time compared to the other physiology sensors, with the complex data analysis tradeoff. The concept of Non-Invasive BCI is a relatively new field of research and it is not until recent years that equipment has made it possible to use online signals produced by the human brain as a vehicle for HCI. Non-Invasive defines that there is no other discomfort to the user than that of wearing the sensor, in other words, there in no injury of the human body involved.

A need arose to combine psychophysiology modalities together, to enrich the interaction experience, and allow further information to be extracted from the data. But this proposition caries with itself a whole new set of problems. Combining modalities became the goal of a multimodal interaction, where ECG, GSR, EMG, temperature, body gesture and many other modalities now share the same weight of making interaction richer and more immersive. Such interaction communicates emotions, or rather emotional states, to the computer to recognize, interpret and communicate emotions back, in the attempt to create believable artificial intelligence (Picard, 2000). Picard further states that we can make user interaction more natural by recognizing the emotional state of the user and exhibiting sympathy and emotional intelligence, which is a trait of intelligent
beings. This interdisciplinary thinking gave birth to a new field of science called *Affective computing*.

*Affective computing* is defined as “computing that relates to, arises from, or deliberately influences emotions” (Picard, 2000, p. 3). In this regard, Picard and colleagues stressed the importance of software intelligence having the ability to sense and respond to the users’ affective states based on physiological measurements (Picard et al., 2001). Moreover, previous research used physiological measures to investigate how interacting with information technology can induce considerable levels of stress in the users (Riedl et al., 2012; Zhai and Barreto, 2006). A core concept in affective computing and also in neuroergonomics is *biofeedback*. Biofeedback aims at displaying physiological parameters such as heart rate or skin conductance acoustically or visually (Lehrer et al., 2000; Nacke et al., 2011; Ouwerkerk et al., 2008a). By this illustration, the subjects get direct feedback on their own visceral processes; processes which are for most subjects usually almost unperceivable (Crone et al., 2004). Biofeedback can increase the user’s attention to his or her emotional state, which in turn can improve *interoception*, i.e. the conscious awareness of one’s own physiological processes (Wang et al., 2010). Previous research has shown that biofeedback can help to successfully reduce heart diseases and depression (Lehrer et al., 2000). In the clinical domain, serious games with biofeedback are brought into use to help young doctors reduce their stress level (Wang et al., 2010). *Physiological computing* represents a category of affective computing that incorporates real-time software adaption to the psychophysiological activity of the user (Fairclough, 2009). The main goal of this is to build a computer that responds to user emotion, cognition and motivation.

*Emotion regulation* can be described by the process model of (Gross, 1998b) which is widely known and acknowledged in the field of emotion regulation strategies. It relies on the assumption that emotions are generated in an emotion generative process. A broad distinction which the author draws is the one between antecedent-focused and response-focused emotion regulation strategies, as depicted on Figure 4. Antecedent emotion regulation strategies apply while the emotion is still unfolding and has not reached its peak. Regulation is influenced by two broad categories of effects: on-task attentional interest and off-task attentional distractors. On-task interest is stronger if the task e.g. is important, if it is intrinsically motivating, rewarded for completion, etc. Off-task distractors are things that make it harder to maintain focus on a desired task, such as disturbing noises, people that interrupt and distractions of an affective character such as sadness or anger. People tend to use one of two main, broadly defined, strategies to deal with emotions emerging when facing difficult and stressful tasks (Wallace et al., 2009). These strategies are *cognitive reappraisal* and *suppression*.
Gross (1998b) defines cognitive reappraisal and suppression as following. Cognitive reappraisal is a type of cognitive change defined as "construing a potentially emotion-eliciting situation in non-emotional terms", the process of constructing a personal interpretation of a situation. Expressive suppression (henceforth referred to just as suppression), is a type of response modulation defined as "inhibiting ongoing emotion-expressive behavior", the process of deliberately trying to stop thinking about certain thoughts. Reappraisal generally reduces emotional experience and behavior expression, and has no impact on memory. Suppression, on the other hand, generally reduces behavior expression, but emotional experience is unaffected and memory is impaired (Fenton-O’Creevy et al., 2012b). They found that reappraisal tendencies correlated with better performance and suppression with worse performance. They explained their results by stating that the cognitive load was higher for suppression than from reappraisal, since reappraisal happens before the onset of an emotion, rendering the emotional response lower and thus creating less off-task attentional distraction. In line with the description in the previous paragraph, suppressors tend to constantly push down emotions, ignoring the fact that they exist and are continuously affecting them. On the other hand, reappraisers tend to positively re-evaluate situations. Both emotion regulation strategies take up cognitive resources (Wallace et al., 2009). However, the authors also state that suppressing emotions generally takes up more cognitive resources in comparison to the reappraisal strategy when encountering undesired emotions. Hence, emotional suppression can eventually take up so much cognitive resources that it can reduce one’s performance in decision tasks compared to the strategy of emotional reappraisal. In fact, emotions and their regulation can be both consciously and subconsciously processed (Gross, 2009; Williams and Bargh, 2009). Following the concept of emotion regulation, emotions can determine and alter the way subjects perceive a certain situation and thereby influence the way they react to it. Thus, emotions act as response tendencies: they suggest distinctive responses to a certain situation, which an individual may or may not follow (Gross, 1998b). In this sense, emotions "reflect the status of one’s ongoing adjustment to constantly changing environmental demands" (Thayer and Lane, 2009, p. 85). Therefore, an increased emotional awareness is crucial for enhanced emotion regulation. The applied emotion regulation strategies seem to be determined by personality traits and individual psychological capacity (Gross, 1998a).

With respect to financial decisions, emotions are suspected to be the underlying driving force for a whole group of biases, such as the disposition effect (Weber and Camerer, 1998), loss aversion (Heilman et al., 2010; Sütterlin et al., 2011), and the phenomenon of auction fever (Adam et al., 2011; Ku et al., 2005). The fact that emotions have an impact on financial decision making became evident in a variety of studies (Adam et al., 2013; Goldstein, 1996; Riedl et al., 2010). It
is important to note, however, that most of these studies refer to contextualized situations and "the context can influence how we process stimuli that may have affective properties" (Rolls and Grabenhorst, 2008, p. 230). In other words, affective processes are context dependent and at this stage it is not clear what in general the emotional ingredients are that guide optimal financial decision making. The emerging field of behavioral finance gives broad evidence that not only financial amateurs, but also financial professional traders and investors suffer from strong decision-making biases (Shefrin and Statman, 1985). Especially periods of high stress and high market volatility can impair economic decision-making and hence trading performance (Lo et al., 2005).

There is broad evidence that emotions are a key factor that critically influences human decision-making (Loewenstein, 2000; Adam and Gamer, 2011). Drawing from economic research there is broad evidence that economic decision making can be biased to a considerable extent by levels of high emotionality and arousal (Loewenstein, 2000; Adam et al., 2011). Fenton-O'Creevy et al. (2011) further detected a strong link between traders' ability to regulate their emotions and their financial performance. The authors found that high performing traders have a better perception and awareness of their emotional state. Most interestingly, these traders are also more advanced in regulating their emotions.

Using this new information body, many serious games and applications in the
domain of clinical medicine today, use biofeedback to help people deal with emotional problems such as anxiety, phobia and post-traumatic stress. It is reasonable to expect that a game that gives instantaneous feedback on a player’s emotional state provides a good pedagogical approach, since direct feedback has been shown to be beneficial in task learning situations (Kluger and DeNisi, 1996). Badly designed games are unappealing, demotivating, and are unlikely to generate useful learning in individuals. As Malone (1980) states, a learning environment has to be optimally complex where the learner knows enough to have expectations about what will happen, but where these expectations are sometimes unmet. Providing some kind of performance feedback to clearly explain why are and why are not expectations achieved, motivates learners for deeper understanding. Malone states that "If students are intrinsically motivated to learn something, they may spend more time and effort learning, feel better about what they learn, and use it more in the future". Learning is the most common of these purposes, based upon the belief that learning through gameplay can be an enjoyable experience that stimulates or encourages players to more intensely engage with educational content than traditional learning media, serious games often deliver situated and action-based learning in simulated environments where physical environments are inaccessible, unavailable, too expensive, or dangerous (Ritterfeld et al., 2009; Corti, 2006).

The De framework enables that the information generated by the different design and evaluation activities finds its way back into the development process. It organizes the overall design and evaluation activities of the project, and reports its findings back to the project. This framework encompasses functionality, heuristic and playtesting evaluation. Functionality evaluation during development consists three categories: validation, verification, and future support. Heuristic evaluation (Isbister and Schaffer, 2008) aiming at qualitatively identifying design errors and suggest improvements. Playtesting evaluation evaluates the tool with the target audience. At present, there exists no unified method to design and evaluate learning games, let alone games for financial capability. Furthermore, no well-documented initiatives exist to help us with the design of financial capability games. The skills and knowledge, attitudes and behaviors that consumers need to enhance their financial security and wellbeing are generally referred to as financial capability (Atkinson and McKay, 2007).

2.1 Related Work

There exists a number of serious games in various fields of science, ranging from medicine, psychology, neuroscience, well-being, finance and many more. Here I
CHAPTER 2. BACKGROUND

will present ones most relevant for this thesis.

2.1.1 Serious Games in the Field of Healthcare and Well-being

There are a number of serious games using biofeedback in the field of healthcare and well-being (Kato, 2010), they are mostly concerned with hyperactivity disorder, autism, substance abuse (Wang et al., 2010) and more specific targets as pain, asthma, bladder control, medical education on cancer. These games use various psychophysiology sensors to make the user aware of his/her medical state and provide a clear goal on how to improve it using feedback (Dunwell et al., 2010). Studies performed by Sharry et al. (2003) show that feedback regarding physiological arousal levels can help people to manage their feeling of anxiety as seen in Relax to Win (Sharry et al., 2003) where a game was created to help children who suffered from anxiety problems. By controlling stress levels, players relax more and thus perform better in the game. The level of the relaxation has been measured using 

\textit{Ecg} and \textit{Gsr}. Froment et al. (2009) enabled participants to play the game of Unreal Tournament while being connected to \textit{Ecg}, \textit{Gsr} and body temperature sensors. By feeding the game with this real time emotion information, the game is manipulated with effects such as screen blur and off-aim. The idea of the game is then to reward the player on successfully resolve the emotional issues. In the game Brainball, Hjelm (2003) prove the potential usefulness of \textit{Bci} in gaming by giving advantage to the most relaxed player. The potential benefits of communicating emotional states is clearly shown in FantasyA and SenToy (Höök et al., 2003) where a physical doll is used to express emotions back to the game by a set of body gestures. Players enjoyed this type of game interaction because they could communicate their feelings to the doll, which is the game avatar itself.

2.1.2 Serious Games using Biofeedback

Psychophysiological signals can be seen as an objective measure of the emotional state, as it is hard to manipulate them intentionally. Iancovici et al. (2011) gives an up-to-date overview on (serious) games using biofeedback, where we can see various usage ranging from educational support (Conati and Chabball, 2003; Kato, 2010) to stress/relaxation/concentration training (Froment et al., 2009; Sharry et al., 2003; Wang et al., 2011). This overview describes serious games in the field of financial decision-making and healthcare. The motivation for serious games presented in this thesis arose from the discrepancy in number
and research in serious games using biofeedback at the time of writing this thesis, between financial decision-making and the healthcare field. This overview marks the former one as lagging behind. Nevertheless, there has been progress in the context of financial decision making; the electrics company Philips together with the Dutch bank ABN AMRO developed a biofeedback device for retail investors. The "Rationalizer" is a device that "acts as a kind of emotion mirror in which the user sees reflected the intensity of his feelings in form of dynamic lighting patterns (Djajadiningrat et al., 2009, p. 39).

2.1.3 Serious Games in the Context of Financial Decision Making

There has also been progress regarding serious games in the context of financial decision making. These games may simulate real life financial situations players will eventually find themselves in, such as the serious game where the player has to make various financial decisions to gather enough money so they can retire and "win" (Jones and Chang, 2011) and "Darwin: Survival of the Fittest" where the player is thought options trading in the trading pit (Michael and Chen, 2005, p. 151); or simulating business and stock market trading, such as a computer-based serious business game where teams of players make various decisions regarding the product manufacture operations of a factory and play the stock market trading company shares, where the team with the highest number of assets is declared the winner after the final period (Hartman and Galati, 2000). Also noteworthy to mention are the games of "Bankloan" and "Supra" where six players take the roles of representatives of three banks/companies seeking to trade loans and three supermarket buyers/sales-men each trading three products respectively (Abt, 2002, p. 101). These articles report students playing these games with enthusiasm as they were used in the curriculum; on the other hand, no measurable quantitative data has been reported on meeting GBL objectives.

In recent years, the advances in cognitive neuroscience have also gained increasing interest from Information Systems (IS) researchers (Dimoka and Banker, 2012; Dimoka et al., 2011; Riedl et al., 2009; vom Brocke et al., 2012). As outlined by Dimoka et al. (2011) and colleagues, the nascent field of NeuroIS is "drawing upon the theories, methods, and tools offered by cognitive neuroscience. This also includes psychophysiological tools such as ECG and GSR measurement (Riedl et al., 2009). Research in the field of NeuroIS has the potential to provide long overdue insight into the decision making processes of users interacting with information technology, where the possibility to map the underlying neural emotional mechanism and activation of ANS exists. For instance, Riedl and colleagues showed that deciding whether to trust an avatar induces less intense neurobio-
logical processes than deciding whether to trust an actual person (Riedl et al., 2011). In a different study, Riedl and colleagues reported that there are gender differences with respect to the activity of specific brain areas when deciding on the trustworthiness of eBay offers (Riedl et al., 2010). Based on heart rate measurements, Adam and colleagues showed that in electronic Dutch auctions the impact of time pressure on final prices is mediated by arousal (Adam et al., 2013).

2.1.4 Summary

As we could see, there exists a number of serious game applications in various fields, of which the ones set in the financial context is the most relevant to this thesis. The articles report implementation description and results, but at present, there exists no unified approach to game design - let alone to serious game design. Nor are there any standardized methods to evaluate games, with serious games posing a particular challenge here. This vagueness of design methods is problematic of scientific game investigation, since it leaves the description space of game features on loose foundations. The ongoing work in the creation of game taxonomies, what Zagal et al. (2005) refer to as game ontology (Caillois, 2001; Aarseth et al., 2003; Björk and Holopainen, 2005; Lindley and Sennersten, 2008) feeds into the design and evaluation methods proposed in this thesis. There is no widely accepted exhaustive consensual taxonomy of game design features for games created to achieve specific affects beyond entertainment (i.e. serious games). Therefore, construct validation requires the identification of a taxonomy that captures design features having an impact upon affects of interest. To ensure that design and evaluation methods proposed are in fact relevant in the context of financial capability, we use the Adult Financial Capability Framework BSA and FSA (2002) developed by Financial Services Authority (FSA) and the Basic Skills Agency (BSA), to develop the different measurement scales of the tool. The FSA/BSA framework, which proposes a set of key skills and competences that characterizes a financially capable person, has guided the design of numerous financial education initiatives.
Chapter 3

Approach

3.1 Aim and Scope

This thesis aims to investigate the process of designing and developing serious games set in the financial context and using psychophysiological methods for training emotion regulation. At present, there exists no unified method to design and evaluate serious games, let alone serious games using psychophysiological methods set in the context of financial capability. By providing a framework for designing and evaluating such serious game applications, the processes and the applications can be scientifically validated. Moreover, combining fundamental theories with different psychophysiological methods can help to recognize and map emotions in a more general and straightforward way, providing a method of incorporating them inside of such serious game applications. Such coupling can support the effects of learning and practicing emotion regulation. Following this conjecture, the scope of this thesis is limited to serious games for emotion regulation training using psychophysiological methods to communicate user’s affective information.
3.2 RQs

The affective information provided by different psychophysiological modalities are diverse, each modality provides a slightly different window to the activation of the ANS. Because of this and the information overlap between modalities, selection of the best modality for a certain problem is not a trivial issue. Regarding this, EEG has been given a separate focus from the rest of the peripheral psychophysiological modalities. For the problem of learning and practicing emotion regulation, serious games have been used as a learning environment. To design, develop and evaluate such serious game application, practices, methods and metrics have been used within the context of The Design and Evaluation (DE) framework. Furthermore, effects of learning emotion regulation have been evaluated. Knowledge from fundamental theories together with different psychophysiological methods allows incorporating mapped emotions inside of serious game applications. Following this, the research questions explored in this thesis are:

**RQ I.** What application possibilities in serious games are there for EEG and other psychophysiological modalities (i.e. ECG, GSR) defined by limitations on available information and response time those modalities provide?

In his book, Andreassi (2007) states that the relation between EEG and human behavior of decision-making plays a prominent role in modern science, since brain processes are the control organ of behavior. However, other physiological modalities (i.e. ECG and GSR explored here) provide important insights into behaviors that are not available through the study of brain activity. With these tradeoffs, the focus on EEG has been explored separately from other physiological modalities mentioned. This question explores metrics of available information that a certain modality can provide, response times needed until we are able to observe a response and application areas where these can be used to enrich the interaction between the application and the user. These emerging differences between EEG and other physiological modalities (i.e. ECG, GSR) are investigated in the Chapter 4 and Chapter 7.

**RQ II.** What is the applicability of practices, methods and metrics proposed to evaluate: serious games (set in the financial context), its design process and design objectives?

Designing serious games is a complex process involving collaborative cooperation between people of various fields. The process has to produce a serious game artifact that meets the design requirements. Furthermore, such an artifact has to fulfill the pedagogical role it has been designed for, or another design iteration has to be made using the methods of design and development to get closer to that role. A need for a framework of practices, methods and metrics emerged.
Proposed practices, methods and metrics are explored in Chapter 5, Chapter 6, Chapter 7 and Chapter 8.

**RQ III. What are the effects of learning, rewarding and practicing emotion regulation measured on the performance in pedagogical approach of serious games?**

Upon the occurrence of emotions, making people aware of those emerging emotions is the first step towards learning emotion regulation. One pedagogical approach is to use serious games to provide an environment where one can learn and practice emotion regulation. In such pedagogical platform, psychophysiological methods have to be implemented as methods of biofeedback to support learning and reward emotion regulation. Such a tool has to match learning objectives set during its design, which have to be evaluated. The effects of learning, rewarding and practicing emotion regulation are explored in Chapter 7 and Chapter 8.

**RQ IV. Is it possible to map the activation of Ans during an emotionally arousing task which is also cognitively demanding?**

Emerging emotions have their neural correlates which unfold in the hard wiring of the human brain. No emotion theory could be explained on the whole if one does not introduce the neuroscientific aspect of their emergence and unfolding. The activation of Ans will be explored in the Chapter 9.

### 3.3 Methodology

The research approach applied in this thesis is based on both quantitative and qualitative methods. These methods include true experiment and evaluation research.

The true experiment constitutes a quantitative approach "to test the impact of a treatment (or an intervention) on an outcome" (Creswell, 2008). This requires that factors affecting the experiment can be controlled and can consequently be called a controlled experiment. True experiments use random assignment of study units (participants) to the experimental treatment or to a comparison group, to ensure that the study units do not affect the outcome instead of the treatment (Kampenes et al., 2009), in the manner of randomized control trial (RCT). Repeated measures design was also used where the same participant was tested under two or more experimental treatments or conditions (Robson, 2002). Exploratory data analysis, or explorative research, is used to investigate little-understood problems, visualize data, and develop question and hypothesizes used
in confirmatory data analysis (CDA) methods (Robson, 2002). True experiments are CDA methods focusing on the testing of the hypothesis.

Robson (2002) states that a questionnaire is a common qualitative method giving a view on the aspects hard to quantitatively measure. But even as such, it provides quantitative data of some qualitative (subjective) enquiry. Securing high degree of involvement by participants is aimed for by administering a self-completion questionnaire immediately after the treatment or the experiment.

Evaluation research explores the effect or effectiveness of some innovation, service, intervention and approach, using a combination of quantitative and qualitative methods. Evaluation, as such, does not necessarily qualify as a research, but characterizes as one when carried out using principled, systematic approach (Robson, 2002). The most important approach to evaluation research is making comparisons between groups, together with other aspects, such as evaluating whether or not the intervention meets the goals and the needs of the users.

### 3.4 Threats

Threats to validity can be divided into four main groups: external, internal, construction and conclusion, with each group containing several threats (Creswell, 2008; Wohlin et al., 2003).

External validity threats concern generalizability and ecological validity of the result. Since we are evaluating a real-world variable in the designed lab environment, we might miss on controlling all the aspects of that variable appearing in the real world. This would render the results of the experiment fault, if inferred on the real-world setting (Wohlin et al., 2003). Tighter control of the experiment might help resolve this issue, but we are still left with the psychological pressure of being measured upon the participant, where he/she is influenced by the presence of the sensor equipment, leading to a distorted outcome of the experiment (Ouwerkerk et al., 2008b). They further motivate the notion that this threat can be studied further in future by introducing remote measuring methods.

Several possible internal validity threats can be ruled out by using RCT design (Robson, 2002). These include selection of participants, history of experiment, since all groups are randomly selected and tested over the same period. Instrumentation threat has been ruled out by ensuring that the same instruments were used on all of the groups, even less so with repeated measurements design. On the other hand, this design introduces maturation threat, as seen in the order effect where fatigue or practice can influence repeated measures at later point in the experiment. Especially interesting are the carry-over effect where there might
be specific effect on the second treatment, resulting from the particular nature of whatever is done first. Sufficient pre-experimental practice helps mitigate the practice, while counter-balancing the treatments helps mitigate order and carry over effects. With questionnaires, there is the memory threat where participants might forget or alter their remembering of the treatment. This issue is reduced if the questionnaire is administered right after the treatment experiment has been carried out.

Construction validity threats are a result of inadequately defined and measured variables (Creswell, 2008; Wohlin et al., 2003). Both quantitative and qualitative variables measured in these publications are, within the domain, standardized and accepted measurements, with long history of scientific use and validation.

Conclusion validity threats concern inaccurately drawn conclusions from the data, referred to as statistical conclusion validity (Creswell, 2008; Wohlin et al., 2003). Through the publications, standardized statistical tests are used to mitigate this issue.

Apart from the validity threats listed, there is another aspect to consider, ethical. Ethical considerations are an important part of any research involving people (Robson, 2002). There is almost always the intention or possibility of change associated with the study. During the experiments participants were purposely faced with situations of stress and anxiety, so careful handling of the situation has to be practiced where participants were informed that they are free to stop and leave the experiment at any time. Participants have to be well informed about the nature of the experiment, up to a point of compromising the experiment. The data was securely stored, according to the ethics and law regulation. These considerations are important aspects of participants’ right to privacy, dignity and self-determination. Moreover, all experiments have to have approval from the ethical committee body revising the experiments in accordance to these considerations.

3.5 Contributions

The following chapter outlines the research contributions that have been published in the articles that constitute Chapter 4 through Chapter 9.

Chapter 4 titled *The Future of Brain-Computer Interface for Games and Interaction Design* presents the potential of BCI paradigm as a non-invasive technology with minimal discomfort, where it has been successfully implemented and tested in various interactive applications and games up to date. Presented and
discussed are potential application areas and uses of modern Brain-Computer Interface (BCI) technology on the example of an EPOC device. Discussion has been divided into two subgroups namely, game design and interaction design, hence the future of these research areas in regards to such technology has been discussed. The contribution of this article is a review on interactive applications and games using BCI technology, in the attempt to present the current State of the Art and infer the future direction. This chapter contributes to RQ I by presenting advantages and disadvantages in using EEG compared to other peripheral psychophysiological modalities. For this publication, the thesis author was the main driver for the investigation, which comprised of reviewing literature and writing the main body of the article. Some insights and additions to the text were done in cooperation with the co-author. This publication was presented at the Fun and Games 2010 conference workshop, Leuven, Belgium, 15-17 September 2010.

Chapter 5 titled Evaluating Games Designed to Improve Financial Capability presents a design and evaluation toolkit capable of providing quantitative results for interactive applications and serious games presented in the Chapter 1, set in the financial context. A multi-level approach to evaluation of serious games for financial capability is presented in this chapter. The approach has been implemented as a toolkit in the context of xDelia, a collaborative project on game-based learning with a focus on emotions in financial decision making. The toolkit has been developed as part of a larger design an evaluation framework for the project. Four facets for financial capability games are targeted by the evaluation: game design, financial capability, behavior change, and learning with technology. The contribution of this article is an investigation on the process and metrics for designing and evaluating serious games set in the financial context. This chapter gives an answer to RQ II, in that it proposes practices, methods and metrics for designing and evaluating serious games set in the financial context. For this publication, the thesis author was not the main driver, but he was involved in the investigation, which comprised of reviewing literature and writing the article in cooperation with the other authors. This publication was presented at the 9th European Conference on eLearning (ECEL) 2010 conference, Porto, Portugal, 4-5 November 2010.

Chapter 6 titled A Serious Game Using Physiological Interfaces for Emotion Regulation Training in the Context of Financial Decision-Making is in part a continuation of the work done in Chapter 5. In this article the Design and Evaluation toolkit has been used to develop a serious game set in the financial context, aiming at helping decision-makers improve their emotion regulation. Research on financial decision-making shows that traders and investors with high emotion regulation capabilities perform better in trading. But how can the others learn
3.5. CONTRIBUTIONS

to regulate their emotions? Learning by doing’ sounds like a straightforward approach. But how can one perform learning by doing’ when there is no feedback? This problem particularly applies to learning emotion regulation, because learners can get practically no feedback on their level of emotion regulation. This research aims at providing a learning environment that can help decision-makers to improve their emotion regulation. The approach is based on a serious game with real-time biofeedback. The game is settled in a financial context and the decision scenario is directly linked to the individual biofeedback of the learner’s heart rate data. More specifically, depending on the learner’s ability to regulate emotions, the decision scenario of the game continuously adjusts and thereby becomes more (or less) difficult. The learner wears an electrocardiogram sensor that transfers the data via bluetooth to the game. The game itself is evaluated at several levels. The contribution of this article is twofold. First, it presents a serious game application using psychophysiological measurements for biofeedback, as discussed in Chapter 4. Second, it uses and evaluates the design and evaluation toolkit presented in Chapter 5, where results show that the developed application met all the requirements of the emotion regulation learning environment. This chapter extends the answer to the RQ II by evaluating practices, methods and metrics proposed in the design and evaluation toolkit; moreover, it evaluates the design process and the metrics for meeting design objectives. For this publication, the thesis author was the main driver for the investigation, which comprised of reviewing literature, developing the game artifact, setting up the experiment design and analyzing the data. The design of the game artifact and writing the article was done in cooperation with the other authors. This publication was presented at the 20th European Conference on Information Systems (ECIS) 2012, Barcelona, Spain, 11-13 June 2012.

Chapter 7 titled A Study of the Effect of Emotion Regulation in a Shooting Game using Bio-feedback from Electroencephalography for Training in Emotion Regulation, presents a serious game as a learning environment for training emotion regulation. The development follows methods and practices proposed in the previous chapters. While emotions are well studied, the regulation of unwanted emotions, and the effects of that regulation, is less well understood. Two broad categories were of interest in the study reported here: cognitive reappraisal, which is associated with less cognitive load and greater well-being, and expressive suppression, which is associated with more cognitive load and less well-being. A biofeedback game, the Aiming Game, was implemented to train the regulation of the emotion component arousal. The effect of emotion regulation on game performance was studied. It was shown that proficient players’ performance correlated positively with expressive suppression, while novice players’ correlated positively with previous experience with shooting games. The results suggest that emotion regulation can be trained using this game, but that expressive suppression would
be the most effective strategy. The contribution of this article is two-fold. First, it presents a serious game application for emotion regulation training. Second, it suggests that two different methods of emotion regulation have a different effect on game performance. This chapter contributes to RQ III by suggesting performance effects on emotion regulation training using two strategies, cognitive reappraisal and suppression. For this publication, the thesis author was not the main driver, but he was involved in the investigation, which comprised of developing the game artifact, setting up the experiment design and writing the article, in cooperation with the other authors. This publication was submitted to the Media Psychology journal.

Chapter 8 titled *Integrating Biosignals into Information Systems: A NeuroIS Tool for Improving Emotion Regulation*, is a continuation of work done in Chapter 6 and Chapter 7. In this article the serious game presented in Chapter 6, has been used to improve emotion regulation in the financial decision context. Furthermore, it evaluated the metrics of meeting design objectives proposed in the same chapter. Traders and private investors are aware that emotional processes can have material consequences on their financial decision performance. However, typical learning approaches for de-biasing fail to overcome emotionally driven financial dispositions; mostly because of subjects’ limited capacity for self-monitoring (Biais et al., 2005, p. 308). Our research aims at improving decision makers’ performance by (i) boosting their awareness to their emotional state and (ii) improving their skills for effective emotion regulation. To that end we design, implement and evaluate a serious game based NeuroIS tool that continuously displays the player’s individual emotional state, via biofeedback, and adapts the difficulty of the decision environment to this emotional state. The design artifact is then evaluated in two laboratory experiments. Taken together, this study demonstrates how to improve financial decision making by integrating tools from cognitive neuroscience in IT artifacts. The contribution of this article is twofold. First, it evaluates the serious game application presented in Chapter 6, on meeting design objectives. Second, it presents a learning environment where financial decision-makers can practice emotion regulation. This chapter extends the answer to RQ II, by evaluating metrics for meeting design objectives; moreover, it contributes to the RQ III, in that it shows the effects on decision performance by learning, rewarding and practicing emotion regulation. For this publication, the thesis author was not the main driver, but he was involved in the investigation, which comprised of reviewing literature, developing the game artifact, setting up the experiment design, data analysis and writing the article, in cooperation with the other authors. This publication was published in the Journal of Management Information Systems (JMIS).

Chapter 9 titled *Identifying the Moment when Demanding Cognitive Load*
Overshadows the Arousal Effect on Pupil Dilation During a Stressful Task, presents an approach to observe sympathetic and parasympathetic pathway activation and dominance over each other, using pupil dilation and heart rate as a measure of arousal. There is a need to develop applications tailored for individual needs and skills. Such applications would need to know when the individual user would perform optimally and when he/she would be overwhelmed, such that the task could be tailored individually for that specific user. This research explores the moment when demanding cognitive load overshadows the arousal effect on pupil dilation during a stressful task. To identify this exact moment, Pupil dilation (Pd) and electrocardiogram (Ecg) data were collected from participants playing a serious game as a dynamic learning environment displaying biofeedback based on Ecg measurements; furthermore, the Pd data was analyzed in relation to the heart rate (Hr) data during this stressful task. Peak of sympathetic nervous system activity was identified in the in the Pd data; thus, this data coupling enabled observation of sympathetic and parasympathetic pathway activation and dominance over each other. The contribution of this article is the method capable of determining the activation and dominance of sympathetic and parasympathetic pathway; moreover, to identify the moment when cognitive load overshadows the arousal effect. This chapter answers RQ IV, in that it provides on method for mapping the activation of ANS during an emotionally arousing and cognitively demanding task based on the activation of sympathetic and parasympathetic pathway measured using pupil dilation and heart rate. For this publication, the thesis author was the main driver for the investigation, which comprised of reviewing literature, developing the game artifact, analyzing the data and writing the main body of the article. Setting up the experiment design and some additions to the text were done in cooperation with the co-author. This publication was submitted to the Applied Psychophysiology and Biofeedback journal and is still under review.

<table>
<thead>
<tr>
<th>Chapter 4</th>
<th>Chapter 5</th>
<th>Chapter 6</th>
<th>Chapter 7</th>
<th>Chapter 8</th>
<th>Chapter 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>RQ I</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>RQ II</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RQ III</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RQ IV</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: Interconnection between chapters and research questions
3.6 Discussion

The techniques investigated in this thesis concern the design and development of affective applications. More specifically, they explore serious games as interactive applications using psychophysiological methods to infer affective information from their users, as described in Chapter 4 through Chapter 6. Furthermore, these techniques investigate the effect of such applications on training emotion regulation, as described in Chapter 7 and Chapter 8. Initially, methods and practices for designing serious games using psychophysiology methods were non-existent and the main premise was to bring into the knowledge from other domains of game design and affective computing. Using these methods for the first time in this context required extensive evaluation both of the methods and developed applications. As presented in Chapter 4, there are quite a few affective applications available, but no design method or evaluation of those applications was reported. Narrower search focusing on the financial context identified the lack of applications and the mentioned methods, as discussed in Chapter 6. Following these notions, The Design and Evaluation framework has been proposed in Chapter 5 combining methods described above. It has been successfully evaluated, both on the design and evaluation process in Chapter 6, and meeting design goals in Chapter 8.

Using psychophysiology methods has been a part of clinical and psychology research for a long time, as discussed in Chapter 6, but only using them online can we consider building affective applications explored in the domain of affective computing. Even though there was an abundance of information and methods from the previously mentioned fields which fed into the affective computing, it is only when using those methods online that new problems and questions arose alongside already known ones. Emotions are still as mysterious in science as they are in real life and the methods of detecting them are as unstable as the constructs they are trying to detect (Russell, 2003). More research is required as this area becomes acknowledged.

Emotion regulation is a technique that helps improve performance on a certain task by controlling the emerging emotions. First step in practicing emotion regulation is to gain awareness of those emerging emotions, so it makes sense to develop tools discussed in this thesis. The Auction game in Chapter 5 and The Aiming game in Chapter 7 were proposed and evaluated, with measurable improvements discussed in Chapter 9 and Chapter 7 respectively. To design a learning environment that would present a user with emerging emotions in order to make him aware, Furthermore, to offer a possibility of training their regulation using the methods of biofeedback, such application should be able to detect such emotions. Two strategies, cognitive reappraisal and suppression, yield different
performance results in a complex task of the Auction Game where cognitive reappraisal is shown to be a more suitable one (Chapter 8), and on an easy task of the Aiming Game where suppression is shown to be more suitable (Chapter 7). A synchronous synergy between the user, affective application, psychophysiological sensors and methods is needed for such a goal to be successful.

Psychophysiology offers a view on bodily reactions to certain emotions, but it is not until we dive deeper into the neural level of the ANS that we discover the complicated and finely tuned mechanism of emotions, orchestrated by the human brain, as presented in Chapter 9. Performance on the task and the affect have two different ANS pathways competing over each other for dominance, parasympathetic and sympathetic pathways. These implications have to be taken into account when designing affective applications for learning emotion regulation, sporting a cognitively demanding task. Only by understanding the underlying neural mechanism of emotions can we truly begin to make general inferences on the observable effects of the bodily psychophysiology.

3.7 Conclusion

RQ I is concerned with the different psychophysiological modalities used in affective applications across the fields of serious games and interaction design. Chapter 4 investigated available information, response time and application areas for the common modalities of EEG, ECG, GSR and some other physiological measures. The review results discovered a growing field of affective applications using mentioned sensors and gave motivation for the possibility of implementing a serious game using psychophysiological modalities online.

RQ II concerned the process and the evaluation of designing and developing serious games using physiological methods in the financial context. In Chapter 5, a Design and Evaluation framework was proposed, feeding design and development practices from the area of game design and affective computing into this new application area. Chapter 6 evaluated methods proposed and put forth an serious game application which fulfilled its design requirements and as described in the Chapter 8, successfully met the design objectives of emotion regulation training in the financial context. Chapter 7 used a different approach by successfully presenting a more general emotion regulation training environment developed with the same methods previously discussed. The results suggest that with careful evaluation, it is possible to design evaluative methods and metrics for affective serious games as learning platforms in emotion regulation training.

RQ III is concerned with the observable effects on the task performance by
practicing emotion regulation with the help of serious game as a learning environment. In Chapter 7, two different strategies for emotion regulation were explored, cognitive reappraisal and suppression, by using a serious game where two groups of participants played it using mentioned strategies. Chapter 8 investigated the effect of practicing emotion arousal on a serious game presenting a stressful task, where groups with and without a helpful learning environment were measured in terms of performance. The results of these experiments show the benefits on decision making performance by using an emotion regulation learning environment. Furthermore, they propose that two different strategies have their benefits in different task contexts.

RQ IV concerned the activation of two different pathways of the ANS on an emotionally arousing task. Chapter 9 explores the underlying mechanism of emotions with the focus on sympathetic and parasympathetic pathway activation of the ANS. Using an emotionally arousing task, physiological measurements of pupil dilation and heart rate were measured to map this activation. For a single individual, the results map the activation and identify the optimal arousal level where arousal effect on the physiology is shadowed by the cognitive demands of the task, thus observing the activation of a sympathetic and parasympathetic pathway. Furthermore, Chapter 9 presents the average optimal arousal across all of the participants, as the goal of the environments which are not dynamic.

The aim of this thesis, guided by research questions presented in Section 3.2, concerned the design and development of affective applications as serious games using psychophysiology methods to detect user’s emotional state. Presented approach shows that it is possible to design and develop such applications that provide helpful learning environment where decision makers can practice emotion regulation and subsequently improve their performance on financial task, which was chosen as a target focus for this thesis. If we consider that emerging emotions have a huge impact on decision making, this motivates the notion that in the decision context, one would benefit from regulating such emerging emotions. Furthermore, we can reach a conclusion why an emotion regulation learning environment is a beneficial approach for those individuals, where psychophysiology methods provide us with the necessary affective input. In conclusion, by careful usage of evaluated design and development methods, affective applications for training emotion regulation can be developed, with successful results improving decision making in its users.
3.8 Future Work

The affective systems presented in the Chapter 6 through Chapter 8 follow the mainstream design logic which state that affective recognition systems today are highly specialized in using one psychophysiological modality, two at most for the most advanced ones. Even though these systems are used to modify the behavior of the applications using them, their main objective is to model and map emotions in a research setting. This leads to high specialization, narrow approach and no generalization properties. Chapter 9 presents a commonly reported problem of conflicting data incoming from different psychophysiological modalities. It is not clear if the problem emerges because of cognitive side of affect models, or psychophysiological mapping and decisions algorithms.

It would be interesting to develop an online affective recognition system, in an attempt to improve methods of emotion recognition and mapping using psychophysiological methods and equipment, both on the cognitive side of the affect models and on the side of decision algorithms. Such system could communicate affective information into the application that the user is interacting with. The application could then respond to his/her emotion, cognition and motivation, inherently making the interaction more significant and richer for both sides.
Chapter 4

The Future of Brain-Computer Interface for Games and Interaction Design

Abstract

In this paper we discuss the potential application areas and uses of modern Brain-Computer Interface technology such as the EPOC. We divide the discussing into two subgroups namely, Game design and Interaction design and hence discuss the future of these research areas in regards to such technology.
CHAPTER 4. THE FUTURE OF BRAIN-COMPUTER INTERFACE FOR GAMES AND INTERACTION DESIGN

4.1 Introduction

The concept of Brain-Computer Interface (BCI) is a relatively new field of research and it is not until recent years that equipment has made it possible to use the signals produced in the human brain as an actual vehicle for computer interaction. In this paper, we look at what has already been done in the field of BCI applications and then propose ideas on how to further extend that research and knowledge. Here, we focus mainly on the use of the BCI tools EPOC, developed by Emotiv.

The paper is divided into two main topics; EPOC for Game design and EPOC for Interaction Design. We choose this separation because even though Interaction design might incorporate gaming and game design, it also represents a cluster of different research areas with very different purposes than those of game design.

4.2 EPOC for Game Design

Many games and applications today use biofeedback to help people to deal with emotional problems such as anxiety, phobia and post-traumatic stress. Since these types of disorders are all highly related to the brain, it is rather likely that Brain-Computer Interface (BCI) technologies, such as the EPOC, could be very effective at detecting and communicating emotions to such an application.

Studies performed by Sharry et al. (2003) show that feedback regarding physiological arousal levels can help people to manage their feeling of anxiety as seen in Relax to Win (Sharry et al., 2003) where a game was created to help children who suffered from anxiety problems. By controlling stress levels, players relax more and thus perform better in the game. The level of the relaxation has been measured using ECG and GSR which have a lag time between the emotion and physical representation. We would like to introduce the EPOC to this type of game concept and upgrade the interaction between the player and the game. Since the EPOC is connected directly to the brain it does not have the lag times that some of the other sensory technologies have. With the EPOC we can detect early signs of tension and help change behavior before spiral arousal/tension has developed fully.

Using BCI as a tool to enhance interaction between the player and the game is an exciting concept. By feeding the game with EPOC sensory information regarding emotions and brain activity, we can modify the game play to drastically increase the game immersion, allowing people to play the game on a whole new level. Our suggestion is not to exclude other interaction devices or sensors, but to
use the BCI equipment to supplement and enrich the interaction experience, much in the same manner as Froment et al. (2009). Here, participants play the game of Unreal Tournament while being connected to ECG, GSR and body temperature sensors. By feeding the game with this real time emotion information, the game is manipulated with effects such as screen blur and off-aim. The idea of the game is then to reward the player on successfully resolve the emotional issues. These types of concepts can enhance emotional self-control as the users are playing against themselves.

Introducing the wireless EPOC, we might be able to substituting multiple sensors, thus giving the player more freedom to interact with the game. Another aspect of the EPOC is that it can communicate many other things than just emotions and brain activity, such as facial expression and head orientation, all of which can make the interaction richer and more immersive. Also, in the game Brainball, Hjelm (2003) prove the potential usefulness of BCI in gaming by giving advantage to the most relaxed player.

4.3 EPOC for Interaction Design

Computers are oblivious to its users’ emotional states. In order to create believable artificial intelligence, such as HAL in Kubrick’s Space Odyssey, it has to recognize, interpret and communicate back emotions (Picard, 2000). By feeding emotional information back to the computer, we can make it interact with user more naturally, recognizing the emotional state of the user and exhibiting sympathy and emotional intelligence which Picard (2000) states as a part of artificial intelligence. Using EPOC, we can directly communicate emotions, subtle facial expressions and brain activity directly to the computer making it aware of the complete emotional state of the person it is interacting with. The ability to communicate emotions is trait of intelligent beings and as such by using BCI we make our computers intelligent. In this manner, we can make computer pick up on emotional trends and act accordingly, learning and communicating with the user.

Using BCI as a tool to enrich interaction by communicating emotions makes for an interesting future. The potential benefits of communicating emotional states is clearly shown in FantasyA and SenToy (Höök et al., 2003) where a physical doll is used to express emotions back to the game by a set of body gestures. Players enjoyed this type of game interaction because they could communicate their feelings to the doll, which is the game avatar itself. The difficulty players had was the built-in indirect control mechanism between doll and the avatar. Using EPOC, we remove this indirect control variable and put more direct control of
the avatar by using emotion and facial expression recognition. Players can then
directly communicate their emotions to the game, making it less "laggy" and
frustrating. The stance of Pervasive Computing is to create context-aware appli-
cations that can adapt to information collected from the environment. Pervasive
Gaming, however, deals with personalized aspects of how players are feeling at
any given moment. As gaming moves more and more towards pervasiveness, we
believe that the EPOC can communicate the player state as well as the envi-
ronmental state to the application, thus enriches every application that uses this
kind of pervasive interaction.

The EPOC is a rather non-invasive technology with no other discomfort to
the user than that of wearing it on the head. As technology moves forward, so
will this device and the possible feeling of discomfort will most likely be reduced.
Brain-Computer Interface is an interesting field with great potential and it is
very exciting to be a part of such a young and blossoming field of research.
Chapter 5

Evaluating Games Designed to Improve Financial Capability

Abstract

A multi-level approach to evaluation of serious games for financial capability is presented in this poster. The approach has been implemented as a toolkit in the context of xDelia, a collaborative project on game-based learning with a focus on emotions in financial decision making. The toolkit has been developed as part of a larger design an evaluation framework for the project. Four facets for financial capability games are targeted by the evaluation: game design, financial capability, behaviour change, and learning with technology. The development of this toolkit is work in progress. An evaluation exercise is planned with existing financial capability games, where we want to assess the toolkit and refine its design to make it more effective for evaluators to use.
5.1 Introduction

This poster illustrates the multi-level evaluation approach of xDelia\(^1\), a research and technology development project funded by the European Commission under the 7th Framework Programme. xDelia is an interdisciplinary project that brings together experts from the fields of organisational behaviour, neuroeconomics, experimental psychology, sensor systems, experimental economics, cognitive sciences, game research, educational technologies, practice-based learning, financial capability, and investment banking. xDelia exploits new and emerging game and sensor technologies to investigate financial decision-making processes, including the role of emotions in people’s decisions. Based on the insights gained from this research, the project will develop new, technology-enhanced approaches to financial training, with support for non-formal and informal learning in real-world settings.

5.2 The design and evaluation framework

The interdisciplinary character of xDelia, the multitude of interlinked empirical studies and game prototypes, and the international nature of the project’s expert team requires an iterative approach to project implementation that is flexible enough to support the complex design and evaluation processes, and, importantly, to enable team learning and a shared understanding of research objectives, strategies, and activities. To ensure that comprehensive, ongoing evaluation is built into all facets of the project and that evaluation findings feed back into the ongoing development activities, xDelia has developed a Design and Evaluation (DE) framework tailored to the project’s specific needs (Clough et al., 2009).

By focusing on the problem of evaluation and design, the DE framework fulfils a dual role. On the one hand, it acts as a guide for designing effective project interventions — workshops, studies, learning games, and so on — providing structure and support for good practice. On the other hand, it acts as a means by which to reflect on implementation processes and on outcomes, involving the stakeholders as reflective evaluators and feeding the findings back into the project on an ongoing basis.

The role of the DE framework is to structure and coordinate the different project interventions, making sure in particular that the information generated by the different design and evaluation activities finds its way back into the development process. We refer to this as macro-level design and evaluation. Macro-

\(^1\)www.xdelia.org
level De thus organises the overall design and evaluation activities of the project, and reports its findings back to the project. For the more detailed specification of individual design and evaluation processes, the De framework provides a micro-level design and evaluation template. For a given project intervention, micro-level De describes the different design and evaluation activities, together with any resources needed for their implementation. Since xDelia consists of numerous project interventions ranging from workshops to studies and prototypes, the collection of micro-level De processes is in fact a network of interrelated design and evaluation activities, where each concrete De process is structured according to the micro-level De template shown in the centre of the poster.

Certain parts of a micro-level De process may be sufficiently well-structured to allow for a stable representation through conceptual schemata or procedures, and can therefore be implemented as paper-based or software tools. The purpose of this article is to report on the first design iteration of a micro-level De toolkit — a set of tools designated for a common use — for financial capability games.

5.3 A design and evaluation toolkit for financial capability games

The skills and knowledge, attitudes and behaviours that consumers need to enhance their financial security and wellbeing are generally referred to as financial capability (Kempson et al., 2005). The financial capability track of the xDelia project focuses on people’s financial skills, attitudes and behaviours rather than on knowledge, and explores whether and to what extent serious games can be effective in this underexplored area. As a foundation for learning game design in financial capability, we have developed a design and evaluation (De) toolkit that focuses specifically on the psychological determinants of financially capable behaviour. The goal is to streamline the De process, making it more efficient and more transparent, and to facilitate collaboration between domain experts and game designers. At the same time, we want to explore more systematically the possibility offered by game-based technologies to modify people’s behavioural patterns and decision making.

At present, there exists no unified method to design and evaluate learning games, let alone games for financial capability. Also, the idea of changing people’s financial behaviour or target the psychological antecedents of maladaptive behaviour with learning interventions is quite new for policy makers and educators. Hence no well-documented initiatives exist to help us with the design of financial capability games. As a starting point, we chose a few topics that we
thought our toolkit should address: financial capability, game design, learning with technology, and behaviour change. Each of these topics offers a useful and unique perspective on the game.

To speed up the game De process, we have developed a set of spreadsheet-based evaluation tools that draw on existing models, frameworks, and taxonomies in financial capability, game design, learning, and behaviour change. The following sections describe the four different tools that make up the toolkit in more detail.

### 5.3.1 Financial capability De tool

A central aim of most financial education initiatives is to provide individuals with the necessary knowledge and skills to make informed financial decisions and appropriate choices. The financial capability De tool evaluates the extent to which a game targets different financial knowledge and skills. To ensure that these are in fact relevant in the context of financial capability, we use the FSA/BSA Adult Financial Capability Framework (BSA and FSA, 2002) to develop the different measurement scales of the tool. The FSA/BSA framework, which proposes a set of key skills and competences that characterises a financially capable person, has guided the design of numerous financial education initiatives.

The current version of this spreadsheet-based tool groups the evaluation into three core sections — knowledge and understanding, skills and competences, responsibility — and nine sub-sections, which are further divides into 113 specific areas of knowledge, skills, and competences. Because evaluators may have problems to assess a game based on the rather vague short descriptions of the nine sub-sections, we have combined the over one hundred specific areas into a more manageable set of 34 categories. Currently, the tool allows evaluators to indicate which financial capability areas the game targets and how well it does this, and provide a weight for each of the evaluation areas.

### 5.3.2 Behaviour change De tool

While the financial capability De tool primarily covers the more conventional financial skills, knowledge, and competences, what we are really concerned with in xDelia are games that in some ways target psychological and social factors that play a key role in motivating and influencing an individual’s financial judgments and behaviour. Much of the psychological and behavioural research in economics has yet to find its way into developing better learning approaches for financial
5.3. A DESIGN AND EVALUATION TOOLKIT FOR FINANCIAL CAPABILITY GAMES

capability education. Since behaviour change — and here we include changes in antecedents of behaviour such as attitudes or perceived social norms — has been widely researched in health psychology, we turn to that literature to develop this De tool. In particular, we draw on standard behaviour models (Glanz et al., 2008) and on a behaviour change evaluation tool for physical activity web sites (Doshi et al., 2003).

In the current version of the tool, we have adopted the taxonomy of behaviour change strategies proposed by Doshi and colleagues — knowledge, cognitive strategies, behavioural strategies, and emotion-focused strategies — and added antecedent factors of behaviour found in standard models such as the health belief model and the theory of planned behaviour. Examples of such factors are self-efficacy, subjective norm, goal setting, feedback, and emotion regulation. Although each of these factors can in principle be changed through learning or training, in practice this might be difficult to achieve or the effect on actual behaviour might be negligible. In general, because very little is known about effective ways to improve financial behaviour, the tool can only be indicative as to possible positive effects of a behaviour change strategy for an individual's financial capability.

The evaluator uses the tool to rate the presence or absence of the different behaviour change strategies within five different levels of increasing player interaction with the learning game. For instance, a game may provide the player with information about a behaviour change strategy, or it may help a player to find out about her maladaptive cognitions, or it may try to change a particular behaviour through gameplay. The levels of interaction are: information, skills, diagnosis, feedback, and gameplay. The tool calculates a score for the depth of player interaction, and we do not foresee at this point to rate the quality of individual strategies, since this is a complicated task that requires very specific expertise for each case.

5.3.3 Game design De tool

Badly designed games are unappealing, demotivating, and are unlikely to generate useful learning in individuals. At present, there exists no unified approach to game design — let alone to serious game design — that we could use as a basis for this tool. Nor are there any standardised methods to evaluate games, with serious games posing a particular challenge here. Here, we use Schell (2008) design lenses paradigm as a foundation for the game design De tool. This is a broad, practice-based approach to game design, incorporating a set of one hundred different design aspect, such as essential experience, problem solving,
competition, feedback, and so on. Each of these lenses asks a unique set of questions about the design, and we have modified them for the purpose of design evaluation.

Evaluating a learning game by asking focused questions about specific aspects of game design is highly intuitive and has a small learning curve for the evaluator, but still leaves room for expert judgment. Not all of the hundreds of questions of Schell’s lenses are relevant for us at this point, and we have therefore reduced them to a more manageable size of 43 questions organised into eight categories: experience, game, elements, theme, iteration, player, mind, and mechanics.

This is the most complex of the four tools, with several auxiliary sheets to support the evaluation process. Each question, or aspect, is assessed on a five-point Likert scale and individual weights can be assigned to account for the importance of a particular aspect in a given evaluation exercise. Based on the rating and weight, scores are calculated for each of the five categories, and a total score gives an indication of the overall quality or merit of the game design.

5.3.4 Learning with technology De tool

Learning design, technology, content, and outcomes are important features of serious game design and evaluation. We are at a very early stage in terms of learning DE tools, and at this point we have added only one tool for learning-specific design and evaluation in financial capability games: learning with technology. The tool draws directly on Jonassen et al. (2002), who proposed five attributes to characterise constructive learning with technology: active, constructive, intentional, authentic, and cooperative learning. So-called assessment rubrics are used to evaluate the extent to which both technologically mediated learning activities and the environment in which they take place promote meaningful learning in formal learning situations. Clough et al. (2009) adapted these rubrics to learning that takes place outside of the formal setting, and our goal is to modify them for learning with games. In next design iteration we plan to add DE tools for learning design and design evaluation, and for learning outcome evaluation.

In the current version of the tool, evaluators assess the games in terms of the 20 sub-rubrics and along the two or three levels of achievement proposed by Clough. A weighted score is calculated for each of the rubrics, where the weights vary as a function of the level of achievement. A first trial with these rubrics has shown that they will have to be modified in order to be operational and useful for serious game design and evaluation. Also, there are clear links between some of the sub-rubrics and the other tools, and these links need to be determined and incorporated in later versions of the toolkit.
5.4 Conclusions

In this position paper we describe ongoing work on a design and evaluation toolkit for financial capability games as part of the xDelia project. The toolkit itself is developed in accordance with the design and evaluation process described in this article. We have now concluded the design part of the first iteration, where we developed four evaluation tools for financial capability, behaviour change, game design, and learning with technology respectively. The remaining work in this iteration is to evaluate the toolkit, which we will conduct using existing financial capability games. The outcome of this evaluation is recorded in the spreadsheet tools and provides the input to the second design iteration which will adapt and refine the tools. The second iteration will also provide us with an opportunity to extend the learning-related DE tools, especially in terms of learning design and learning outcomes.
CHAPTER 5. EVALUATING GAMES DESIGNED TO IMPROVE FINANCIAL CAPABILITY
Chapter 6

A Serious Game Using Physiological Interfaces for Emotion Regulation Training in the Context of Financial Decision-Making

Abstract

Research on financial decision-making shows that traders and investors with high emotion regulation capabilities perform better in trading. But how can the others learn to regulate their emotions? 'Learning by doing' sounds like a straightforward approach. But how can one perform 'learning by doing' when there is no feedback? This problem particularly applies to learning emotion regulation, because learners can get practically no feedback on their level of emotion regulation. Our research aims at providing a learning environment that can help decision-makers to improve their emotion regulation. The approach is based on a serious game with real-time biofeedback. The game is settled in a financial context and the decision scenario is directly linked to the individual biofeedback of the learner’s heart rate data. More specifically, depending on the learner’s ability to regulate emotions,
the decision scenario of the game continuously adjusts and thereby becomes more (or less) difficult. The learner wears an electrocardiogram sensor that transfers the data via Bluetooth to the game. The game itself is evaluated at several levels.
6.1 Introduction

Serious games are (digital) games used for purposes other than mere entertainment (Susi and Johannesson, 2007). Corti (2006) points out an obvious advantage of serious games in allowing learners to experience situations that are impossible in the real world for reasons of safety, cost, time or logistics. Serious games can have positive impact on the development of a number of different skills (Ellis et al., 2006; Mitchell and Savill-Smith, 2004; Corti, 2006; Squire and Jenkins, 2003; Van Eck, 2006; Rieber, 1996) Having those different skills as defined learned outcomes, one can clearly see why serious games are considered as Game-based Learning (Gbl) applications (Corti, 2006). To achieve the development of new knowledge and skills, game-based learning/serious games need to captivate and engage the end-users for a specific purpose (Corti, 2006). Corti further states that Gbl has the potential of improving training activities and initiatives by virtue of, e.g., its engagement, motivation, role playing, and repeatability (failed strategies etc. can be modified and tried again); thus, lead to a more productive workforce.

Currently, there have been serious games created with the goal of teaching how to better manage financial decisions. These games may simulate real life financial situations players will eventually find themselves in, such as Massively Multiplayer Online Role Playing Game where the player has to make various financial decisions to gather enough money so they can retire and "win" (Jones and Chang, 2011) and "Darwin: Survival of the Fittest" where the player is thought options trading in the trading pit (Michael and Chen, 2005, p. 151); or simulating business and stock market trading, such as a computer-based simulation business game where teams of players make various decisions regarding the product manufacture operations of a plant(s) and play the stock market trading company shares, where the team with the highest number of assets is declared the winner after the final period (Hartman and Galati, 2000). Also noteworthy to mention are the games of "Bankloan" and "Supra" where six players take the roles of representatives of three banks/companies seeking to trade loans and three supermarket buyers/sales-men each trading three products respectively (Abt, 2002, p. 101). These articles report students playing these games with enthusiasm as they were used in the curriculum; on the other hand, no measurable quantitative data has been reported on meeting Gbl objectives.

Classical economic theories and models are usually based on the assumption of market actors being fully rational and favor utility maximization when confronted with economic decisions (Rasmusen, 2003). This way of considering economic decision-making has not only dominated economic literature for decades but has also shaped how humans perceive their economic decisions. In
particular, professional investors and traders are considered to behave perfectly rational. However, the emerging field of behavioral finance gives broad evidence that not only financial amateurs, but also financial professional traders and investors suffer from strong decision-making biases (Shefrin and Statman, 1985). Especially periods of high stress and high market volatility can impair economic decision-making and hence trading performance (Lo et al., 2005).

There is broad evidence that emotions are one key factor that critically influences human decision-making (Loewenstein, 2000; Adam and Gamer, 2011). As will be shown in the next section, emotions do not always impair decision making. They can also have a positive influence and facilitate decisions. Gross (1998b) argues that emotions act as response tendencies and subjects may follow these response tendencies or not. Recent research shows that, the ability of detecting or being aware of one’s emotions and the skills to down-regulate levels of high emotional arousal improves human decision-making (Fenton-OCreevy et al., 2011).

Following this conjecture, the international project xDelia (Xcellence in Decision-making through Enhanced Learning in Immersive Applications, www.xdelia.org) has developed a serious game that aims at improving the player’s emotional awareness and training of his/her ability to down-regulate levels of high emotional arousal by the means of online information system displaying biofeedback based on psychophysiological measurements. As an advantage this system measures reliable emotional arousal in a stressful environment and is not biased by self-perception.

Based on the game, future experimental research can shed more light on the connection of training of emotion regulation and decision-making. Moreover, a bank with high expertise in the private investors sector will test the game as a training tool for real traders and investors in day trading centers.

The remainder of the paper is structured as follows. In Section 2, we will first describe the theoretical background on emotions and decision-making upon which the development of the game is based. Section 3 describes how emotional arousal can be measured externally with the use of psychophysiological measurement technology. We then describe the design of our game approach — which we titled Auction Game — and present and discuss evaluation results in terms of game functionality and usability/playability.
6.2 Emotions, Emotion Regulation and Decision-Making

Drawing from economic research, there is broad evidence that economic decision making can be biased to a considerable extent by levels of high emotionality and arousal (Loewenstein, 2000; Adam et al., 2011). In the context of economic decision making, emotions are usually perceived as inappropriate interfering with the rational best decision and impairing the decision-maker’s ability to take "good" decisions. For instance, the disposition effect, i.e. the tendency for cashing in winning stocks quickly while holding on to losing stocks for too long, is often explained by subjects’ emotional imbalance of how to cope with gains and losses (Shefrin and Statman, 1985). However, Seo and Barrett (2007) discovered in an empirical investigation with traders that emotions may also have positive effects on their stock trading performance. There is hence a bilateral effect of emotions: on the one hand, they are bias-inducing and hence malicious to the decision maker, but on the other hand, they also provide valuable knowledge in representing for example experiences one has gained in the past (Bechara and Damasio, 2005; Astor et al., 2011). Emotions can help evaluating situations instantly or processing informational overload, when one has to come to quick decisions.

As mentioned above, emotions also affect decision making in professional settings, such as trading, which originally was believed to be a purely rational act. Several studies give evidence that professional traders are tremendously influenced by their emotions. Fenton-O’Creevy et al. (2011) interviewed a set of traders in detail and reported that periods of losses were often accompanied by very risk-averse behavior and cautiousness. However, major gains often resulted in high confidence and headless behavior. These emotional states are often accompanied by high emotional arousal.

Russell (1980) generally classified emotions by their independent components arousal and valence. Thereby, arousal represents the level of excitement whereas valence defines whether the current emotional state is positive or negative (visualized on Figure 1). Following this notion when measuring emotions, one is actually measuring a combination of valence and arousal. A reliable measure which shows different kinds of variation depending on the kind of emotional stimulus is the heart rate (Anttonen, 2005; Vrana et al., 1986; Leng et al., 2007). Furthermore, since levels of high arousal can be accompanied by positive as well as negative emotions, arousal remains as the primary attribute of interest in our study. In the scope of our game approach we use heart rate as a proxy for emotional arousal, which will be described later. The continuous measurement of heart rate helps to improve the understanding of the emotional processing in economic decision making.
Fenton-OCreevy et al. (2011) further detected a strong link between traders’ ability to regulate their emotions and their financial performance. The authors found that high performing traders have a better perception and awareness of their emotional state. Most interestingly, these traders are also more advanced in regulating their emotions. While less experienced traders usually try to avoid aversive emotions, the more experienced traders had actually learnt to cope with their emotions. Consequently, the more experienced traders were able to identify and discriminate their emotions in a more sophisticated way. Thus, there are interpersonal differences considering the experience, the awareness, and the ability to regulate emotions, which in turn inhibit or facilitate decision making performance.

Figure 6.1: Emotions in the valence-arousal space
6.2. EMOTIONS, EMOTION REGULATION AND DECISION-MAKING

Emotion regulation can be described by the process model of Gross (1998b) which is widely known and acknowledged in the field of emotion regulation strategies. It relies on the assumption that emotions are generated in an emotion generative process. A broad distinction which the author draws is the one between antecedent-focused and response-focused emotion regulation strategies. Antecedent emotion regulation strategies apply while the emotion is still unfolding and has not reached its peak. An example for emotional reappraisal would be a shy student in a class. Now, emotional reappraisal could result therein that s/he considers the school class as a good opportunity to train raising his/her hand and answering questions. Hence s/he constructs a potentially emotion-eliciting situation in nonemotional terms. Response-focused emotion regulation on the other hand tries to aim at altering and controlling the experiential, behavioral and physiological response to the fully established emotion. An example for such behavior could be the shy person in a school class which might try inhibiting ongoing emotion-expressive behavior and disguise them with e.g. insubordination.

People tend to use one of two main, broadly defined, strategies to deal with emotions emerging when facing difficult and stressful tasks (Wallace et al., 2009). These strategies are reappraisal and suppression. In line with the description in the previous paragraph, suppressors tend to constantly push down emotions, ignoring the fact that they exist and are continuously affecting them. On the other hand, reappraisers tend to positively re-evaluate situations. Both emotion regulation strategies take up cognitive resources (Wallace et al., 2009). However, the authors also state that suppressing emotions generally takes up more cognitive resources in comparison to the reappraisal strategy when encountering undesired emotions. Hence, emotional suppression can eventually take up so much cognitive resources that it can reduce one’s performance in decision tasks compared to the strategy of emotional reappraisal.

Gross and John (2003) designed a questionnaire — the Emotion Regulation Questionnaire (ERQ) — in order to identify suppression/reappraisal strategy tendencies used by individuals. It makes specific statements with respect to the emotion regulatory process intended to be measured, such as "I control my emotions by changing the way I think about the situation I’m in."

This section has shown that emotions in the context of financial decisions can be both, bias-inducing, and performance boosting. Better emotion regulation strategies result in better financial performance, whereby the awareness of the emotional state seems to be critical for appropriate evaluation of the decision situation. The developed serious game for our study thus aims at improving emotion regulation, but also to improve the players’ emotional awareness.
6.3 Psychophysiological Measurement of Emotions

In order to make players aware and give sufficient feedback on emotional arousal, it is crucial to apply a method to reliably detect emotional arousal. While subjective measures, such as self-evaluation, always also incorporate potential self-deception, we make use of psychophysiological correlates of emotions. Moreover, psychophysiological signals can be seen as an objective measure of the emotional state as it is hard to manipulate them intentionally. Iancovici et al. (2011) gives an up-to-date overview on (serious) games using biofeedback, where we can see various usage ranging from educational support (Conati and Chabбал, 2003; Kato, 2010) to stress/relaxation/concentration training (Froment et al., 2009; Sharry et al., 2003; Wang et al., 2011). There are a number of serious games in the field of healthcare and well-being using biofeedback (Kato, 2010), they are mostly concerned with hyperactivity disorder, autism, substance abuse (Wang et al., 2010) and more specific targets as pain, asthma, bladder control, medical education on cancer. These games use various psychophysiology sensors to make the user aware of his/her medical state and provide a clear goal on how to improve it using feedback (Dunwell et al., 2010). This last point gives the motivation for the game presented in this paper; even more so if we consider that this overview describes the field of finance and financial decision-making as lagging behind in using serious games with biofeedback, in contrast to the healthcare field.

For computation of heart rate (HR), we used the ekgMove sensor developed by Movisens which records electrocardiographic (ECG) signals with high accuracy. In contrast to most other commercially available ECG devices, the sensor is attached to the chest using a flexible belt with dry electrodes. Therefore, it is less obtrusive than other devices and offers a higher wearing comfort. The ECG signal is transmitted via Bluetooth to the xAffect software environment (Schaaff et al., 2011). The software offers a modular framework which allows to process data from various input devices and to transfer the derived values via TCP/IP to other applications like the Auction Game. To get information about the current arousal level of a person, the heart rate is computed from the raw ECG signal. An algorithm to derive the current arousal level from heart rate information is implemented in the xAffect framework. The arousal levels are computed in relation to a baseline period which is recorded before the game starts.
6.4 Game description

6.4.1 Underlying principle

The developed game serves two major goals:

- Improvement of introspection, the examination or observation of one’s own mental and emotional processes, and self-monitoring of physiological arousal and hence personal emotional state.

- Improvement of skills in emotion regulation by elements that reward good emotion regulation and punish poor emotion regulation strategies.

In order to achieve these goals, the game uses a physiological interface detecting online physiologically measured levels of arousal, as a basis for providing emotional feedback (biofeedback); furthermore, the game difficulty is connected to the measured level of arousal. The better the player is able to control and adapt his/her level of arousal, the easier the decision environment is.

The core motivation for the Auction Game is that there is a link between maladaptive financial behavior and poor emotion regulation abilities. Therefore the Auction Game can be considered as an emotion regulation training game in the context of financial decisions.

6.4.2 Game concept and gameplay

The narrative in the Auction Game is purposely simple, since it has to be easy to use the game for students, as well as for investors in day trading centers. The theme of the game is an abstract one depicting sky and clouds, as a supporting environment for down regulating levels of high emotional arousal.

The player is set in the position of a trader where s/he continuously can buy or sell goods, in each round one at a time. The game starts with the introductory screen where the player is presented with the instructions on how to play the game. Here it is also possible to go through the tutorial or just start the game, after which the player’s baseline HR data is recorded. Every further arousal level measure in the game is calculated against this baseline. A previous tutorial explains the principle of the game and slowly guides the player. The tutorial should be played the first time the player gets in contact with the game, but it can be skipped if the player already knows the Auction Game.

Before the start of each round, an offer price and price estimations are calculated with respect to the level of arousal the player is currently at. A round
(see Figure 2) consists of three price estimations presented to a player sequentially, from the three trusted simulated consultants. The clouds are individually presented on screen for a certain amount of time, from one second at starting easy levels and shorter as levels progress. To make the player attentive, clouds appear at random places on the screen. They estimate the goods price in the next round; thus, by calculating the mean of the three prices in the clouds, the player knows the true price of the good/stock in the following round.

After the indication of three price estimations, the player gets the chance to buy or sell the good for the offered price. S/he has to make a decision in a certain amount of time, from two seconds at easy starting levels and shorter as levels progress. To make a decision, the player has to click on the buy or sell button; following this, an audio and video feedback is presented conveying the outcome of the decision.

![Game rounds](image)

**Figure 6.2: Example sketch of game rounds**

Depending on his/her decision the player realizes a gain or a loss. In order to gain money the player has to take the right buy/sell decisions. Every profitable decision will reward the player with a certain amount of money, while a non-profitable one will reduce the player’s earnings and take him further away from the current level’s profit goal. Not taking a decision is the most expensive action, taking a large sum of profits (5 Euros) away from the player. Too many money losses will lead to the end of the game. This limit has been set to 10 Euros below the current level starting point. If the player is quick in calculating the three price
estimations, s/he can easily reach a correct decision. After that, the total money earnings are updated and if the player has reached the current level’s profit goal, a new round begins.

Consider the example in Figure 2. Assume that the price estimations are 92.48 €, 93.31 € and 87.80 € and that the offered price is 95.42 €. By calculating the mean the player realizes that the mean is close to 91 €. In fact, it is 91.20 €. Since the mean price of the estimations is lower than the offered price, the player should select the Sell button and realize a gain of 4.22 €. This task continues until the player has reached an upper bound to move to the next level or the player loses the game due to running out of time or due to bankruptcy.

The player’s level of physiologically measured arousal affects the game difficulty. Before the start of each level, player is informed on how his/her level of arousal will influence the game difficulty. For example, the goal within a round could be to keep the level of emotional arousal as low as possible. As long as the player is able to keep arousal low the game will remain in the relatively easy mode. However, as soon as s/he becomes more aroused due to, for example, anger of an incorrect decision, the arousal bar will move up and the decision task will become more difficult. The level of difficulty increases by increasing the variance on the price estimation signals. While the price signals would normally (without arousal) be 92.48 €, 93.31 € and 87.80 € they could be then 68.22 €, 79.21 € and 126.17 € when the player is unable to keep his/her arousal down, making it more difficult to calculate (or estimate) the mean. The variance of the estimations will get larger the higher the emotional arousal is.

The player’s goal in the Auction Game is to reach the highest level possible. To advance from one level to the next, the player has to make profitable decisions and earn enough money to reach that level’s profit goal. Moreover, s/he has to reach the next level in a limited period of time otherwise the game ends; this time limit is currently set to 4 minutes. In every level the profit goal increases by 30 Euros. As soon as the player earns this amount, a button pops up and s/he can proceed to the next level.

Players who can achieve targeted the emotional arousal level will be rewarded with larger profit/lesser loss money values. On the other hand, undesired emotional arousal values will yield lesser profit and larger loss. While in the first levels the constraint for physiological arousal is to simply down-regulate arousal, in later levels the players have to aim at a specific level of arousal. As the game progresses, as well as arousal gets higher, the game difficulty will change, making it more difficult to take a profitable decision. Moreover, from level to level the tasks to regulate emotions will become more and more difficult as step by step new distracting elements are included. Among those are additional irrelevant
clouds carrying false information, time, distracting images, auditory, and visual constraints (will be described later in detail).

The game is conceptualized such that it has no predetermined ending level, but after level 10 the game becomes extremely hard entering what is called ”The Death Mode”. Hence, the ending time of the single levels and the game can vary, depending on the player’s skills. Optimally the game should run for approximately 25 minutes where players should earn around 200 Euros profit in the game. The player’s skills to earn money in the game are related obviously to (1) his/her abilities to perform calculations under stress, and (2) to the player’s skills to regulate his/her overall emotional arousal state. Independently from the player’s calculation skills, good emotion regulation will help to improve his/her individual game performance.

After the game has ended, the player is presented with the level s/he has reached. Better players earn a higher place on the high score list where they are given a chance to compare money earned result to their previous ones.

The objective of the Auction Game is to train players in performing emotion regulation strategies. By showing the level of arousal the player can gain an awareness of his/her emotional state and the influence of emotions and emotion regulation on decision making. In other words, guide the player towards mindfulness of emotions. By displaying the player’s emotional state as an indicator indicating arousal levels ranging from relaxed to highly aroused, a player has to regulate his/her arousal to minimize the deviation of the estimations and thereby have a better chance to accomplish a higher profit. Indicator levels are dependent on a player’s level of arousal.

6.4.3 Game Mechanics

The Auction Game has been developed in a Unity 3d pro game engine which supports integration of third party APIs. As can be seen in the Figure 3, the Auction Game is played in a 2D environment where price estimations are presented inside of the colored cloud drawings. To depict a sense of progress through the game, every level has a different background picture of the sky. The player can see his/her individual arousal level indicated on the meter in the top right corner, as well as by the color of the clouds (green, yellow and red). The profit goal and total money earned are presented on the meter at the bottom right side of the screen. Decisions can be made by clicking on one of the Buy/Sell buttons presented at the bottom of the screen using a mouse.
6.4. GAME DESCRIPTION

Figure 6.3: Screenshot from the Auction Game
CHAPTER 6. A SERIOUS GAME USING PHYSIOLOGICAL INTERFACES FOR EMOTION REGULATION TRAINING IN THE CONTEXT OF FINANCIAL DECISION-MAKING

Game Difficulty  The game must engage in play all types of players, ranging from experienced gamers to completely inexperienced players since the target group for a serious game may not necessarily be experienced game players. Thus, the game starts slowly introducing distracting elements step by step throughout the levels; moreover, the clouds appearance time (mean calculation) and decision time decreases.

In the Auction Game, different game elements are affected by the player’s arousal level, which will make the game harder. In order to train emotion regulation during the game, it is important that the game is sufficiently challenging in the aspect of emotional arousal control. The arousal affected elements described below are the different ways in which arousal influences gameplay, and are meant to make the game more difficult in different aspects. The further away the player’s current level of arousal is from the target level, the bigger each of the effects will be.

Thus there are two dimensions of variety of difficulty in Auction Game respectively:

- **Game elements not affected by arousal**
  - As soon as the player reaches half of the level goal, the tempo of the background music will be slightly increased. Moreover, one quarter away from the goal music noticeably speeds up to distract the player and thereby suborn him/her to make quick decisions.
  - The speed of cloud appearance increases while the time for decision decreases. As the player progresses through a level, the cloud estimations slightly increase their movement speed; moreover appearance time and decision time slightly decrease to distract the player into making quick decisions in an attempt to make him commit errors.

- **Game elements affected by arousal**
  - Distribution of price signals is dependent on the arousal level of the player. The further s/he is away from the target arousal, the larger will be the spread of estimations. This will make it more difficult to calculate the true price.
  - Distribution of true price in the next period is dependent on the arousal level of the player. Every round a good shifts its true price on the market. The further s/he is away from the desired level of arousal, the larger deviation of the next true price is. This will make true price shift more unexpectedly.
6.5. EVALUATION

— Speed of cloud movement is directly linked to arousal. As the game progresses clouds start moving. The further s/he is away from the desired level of arousal, the faster the movement of clouds becomes. This makes it harder to visually observe the price estimation.

To keep the game interesting, piecewise elements varying cloud estimations are presented through the levels (Table 1). Note that every element adds to all the active ones from previous levels.

<table>
<thead>
<tr>
<th>Level</th>
<th>Elements varying cloud estimations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>Player’s arousal level is presented but it has no effect on the game at all</td>
</tr>
<tr>
<td>2nd</td>
<td>Player’s arousal level affects the game</td>
</tr>
<tr>
<td>3rd</td>
<td>Estimation clouds move simulating the wind</td>
</tr>
<tr>
<td>4th</td>
<td>Estimation clouds become bigger and smaller in sizes</td>
</tr>
<tr>
<td>5th</td>
<td>Estimation clouds are same sized, but fake clouds with text start to appear</td>
</tr>
<tr>
<td>6th</td>
<td>Fake clouds with numbers start to appear</td>
</tr>
<tr>
<td>7th</td>
<td>Estimation clouds become bigger and smaller in sizes again</td>
</tr>
<tr>
<td>8th</td>
<td>Player has to achieve targeted arousal, but fake clouds do not appear in this level</td>
</tr>
<tr>
<td>9th</td>
<td>Fake clouds with numbers start to appear again</td>
</tr>
<tr>
<td>10th</td>
<td>Estimation clouds’ speed and appearance time of the clouds are random</td>
</tr>
<tr>
<td>11th</td>
<td>Entering Death Mode, game speeds up impossibly</td>
</tr>
</tbody>
</table>

Table 6.1: Game level elements varying cloud estimations

6.5 Evaluation

The Auction Game has been thoroughly tested using functionality, heuristics and play testing.

6.5.1 Functionality and Heuristic Evaluation

The functionality evaluation consists of three categories: validation, verification, and future support. The validation was done in collaboration between the partners from xDelia at several face-to-face meeting where the product was demon-
strated. By letting the product owner test it before it was finalized much of the rather abstract discussions became concrete as feedback was communicated.

The verification was done by the development team throughout the development process. By working in an incremental development style the product was constantly tested for errors when new things were added, thus minimizing latent errors. The future support is achieved by designing the game around a modular and dynamic architecture, making it easy to adjust for future studies; moreover, a xAffect as separate component is used for measurement of the arousal allowing for other sensors to be used within the same game.

Development of the Auction Game was followed by a heuristic evaluation (Isbister and Schaffer, 2008) aiming at qualitatively identifying design errors and suggest improvements. A group of expert evaluators reviewed the game using heuristics divided into a set of categories, looking for potential usability and gameplay problems in the prototype. Heuristic evaluation pinpointed several important design issues and reinforced the robustness of the game.

### 6.5.2 Playtesting Evaluation Method

A total of six students volunteered to participate in the Playtesting Evaluation. They were all students of Blekinge Institute of Technology aged between 20 and 32 years old with four of them being male and two female. They reported varying gaming experience.

Before the game, the students were fitted with the Movisens ecgMove HR sensor and given a tutorial session. In order to objectively determine which game elements the players were paying attention to, the game was played through the Tobii T60 eyetracker logging data on different Areas of Interest (AOIs) and recording the whole gaming session on video. The purpose was to be able to tell how important different visual objects (AOIs) on screen are to a player. At the end of each game, each participant was given: an Emotion Regulation Questionnaire (Gross and John, 2003) in order to identify suppression and reappraisal tendencies of individuals; a modified System Usability Scale (Nacke et al., 2010; Brooke, 1996) questionnaire measuring game usability. The questionnaire contained 10 questions whose score was summed up in a single number representing a composite measure of the overall usability of the game being studied; an interview session where participants could openly discuss perceived game speed and difficulty, as well as visual cue elements and any issue they wanted to note.
6.5.3 Playtesting Evaluation Results

The Auction Game scored a mean value of 67.92 in a range from 0 to 100 on modified SUS questionnaire. Thus according to (Tullis and Albert, 2008) where a score of 60 presents a border between poor and average usability, we can conclude that the game fulfills the average game usability.

The game was successfully played up to a hard 8th level by two participants, both of which are high reappraisers, while one had low and the other normal suppression tendencies. They both evaluated the game as manageable and in the interview session reported that they were practicing emotion regulation techniques themselves without being instructed at all. This provides evidence for a good game design of the Auction Game.

Five out of six participants reported that they were not paying attention at all to the arousal meter indicator present at the top-right of the screen. We evaluated this claim on how informed the participants were about their arousal level. They have an option of keeping track of it on the arousal meter indicator during rounds in the whole gaming session. A paired-samples t-test was concluded on the eyetracker data to evaluate the difference in number of gaze observations on marked indicator arousal meter AoI compared to number of rounds taken for each participant. There was a statistical significant difference found with number of rounds (M = 110.17, SD = 95.26) to number of arousal meter observations (M = 16.33, SD = 24.5, \( t(5) = 2.94, p < .05 \)). Thus we can say that participants paid little or no attention on the arousal meter indicator during the whole playing session. Participants reported that the reason for paying little or no attention on the arousal meter indicator was lack of time during fast paced decisions. Most of the participants reported that they were paying attention to the arousal indicated by the color of the cloud estimations, especially when it turned red. This gave evidence to concentrate on making the color of the cloud estimations more distinct, since players are focusing their concentration on them. Further studies should identify how to optimally present the arousal information to the cognitively engaged player during fast paced decisions.

A one-way between-groups analysis of variances was conducted to explore the impact of arousal level on profit in each round. Total number rounds played, 661 rounds, were divided into 5 groups according to the arousal level while decision was made (Group 1: 1[relaxed], Group 2: 2 ... Group 5: 5[highly aroused]). There was a statistically significant difference at the \( p < .05 \) level in profit made each round for the five arousal groups \( [F(4, 656) = 3.566, p < .01] \). The effect size, calculated using eta squared was .02. Post-hoc comparison using Turkey HSD test indicated that the mean score for Group 1 (M = .6328, SD = 2.95) was significantly different from Group 5 (M = -1.369, SD = 3.3). Other groups
CHAPTER 6. A SERIOUS GAME USING PHYSIOLOGICAL INTERFACES FOR EMOTION REGULATION TRAINING IN THE CONTEXT OF FINANCIAL DECISION-MAKING

did not differ significantly. Same has been conducted for the time needed to reach a decision in seconds and there was a statistically significant difference the \( p < .05 \) level \( [F(4, 656) = 5.753, p < .001] \) between Group 5 (\( M = 1.55, SD = .45 \)) and rest of the groups. The effect size, calculated using eta squared was .03. This gives strong evidence supporting a good design of the Auction Game to present a hard challenge and punishment to a player in an undesirable high arousal emotional state.

6.6 Discussion and Conclusion

Evidence shows that emotions impact our decisions, especially in the field of finance. Thus it makes sense to develop a tool to get people aware of this implication as well as to help them in regulating their emotions to reach better financial decisions. A serious game emerged as an appropriate tool in which players get feedback on their emotional arousal, according to their psychophysiological state. This on screen feedback helps subjects to get aware of and to learn how to control their emotional state. The Auction Game is a serious game where a player buys or sells stocks with the objective to train emotion regulation; but also to get them aware of the arising emotions. To support this, achieving a target arousal level will reward the player accordingly, increasing his/her earned profit.

Data from successful participants gives first evidence that the Auction Game is indeed overwhelming and puts players at a highly aroused state where they need to practice emotion regulation techniques to succeed in the game.

We have demonstrated in the Auction Game how one can reward a player achieving a desired arousal level, while at the same time presenting a hard challenge and punishment to a player in an undesirable high arousal emotional state. Through this experience emotion regulation can be learned and practiced using this tool.

For future work it is planned to use the Auction Game in varying contexts: it will be interesting to detect which strategies players apply (e.g. breathing) in order to regulate their level of emotional arousal; Related to this question, we want to examine how effective certain strategies will prove, measured either by self-perception and/or by physiological measures (e.g. phasic heart rate response, heart rate variability), in order to perform well in the game. Moreover we want to find out whether certain emotion regulation strategies (suppression versus reappraisal) result in systematic differences in game performance. Our last and most prominent goal is to evaluate the Auction Game as a learning tool for enhanced emotion regulation, i.e. examining whether extensive playing of the
6.6. DISCUSSION AND CONCLUSION

Auction Game (or another tool following the same paradigm) can systematically improve subjects skills to get aware and control their emotions and whether these skills are transferable to other (financial) tasks, leading to a long lasting shift in decision performance. This future research has to be conducted in order to investigate how successful our approach is in teaching emotion regulation and how well it can be transferred to real life trading. Up to now, we have demonstrated that the Auction Game was successful at reaching its goals as a study tool, as well as a usable game. If we can systematically succeed in this, we can make learning emotion regulation in the context of financial decision making more fun and more effective.
CHAPTER 6. A SERIOUS GAME USING PHYSIOLOGICAL INTERFACES FOR EMOTION REGULATION TRAINING IN THE CONTEXT OF FINANCIAL DECISION-MAKING
Chapter 7

A study of the effect of emotion regulation in a shooting game using bio-feedback from electroencephalography for training in emotion regulation

Abstract

While emotions are well studied, the regulation of unwanted emotions, and the effects of that regulation, are less well understood. Two broad categories were of interest in the study reported here: cognitive reappraisal, which is associated with less cognitive load and greater well-being, and expressive suppression, which is associated with more cognitive load and less well-being. A bio-feedback game, the Aiming Game, was implemented to train the regulation of the emotion compo-
CHAPTER 7. A STUDY OF THE EFFECT OF EMOTION REGULATION IN A SHOOTING GAME USING BIO-FEEDBACK FROM ELECTROENCEPHALOGRAPHY FOR TRAINING IN EMOTION REGULATION

... arousal. The effect of emotion regulation on game performance was studied. It was shown that proficient players’ performance correlated positively with expressive suppression, while novice players’ correlated positively with previous experience with shooting games. The results suggest that emotion regulation can be trained using this game, but that expressive suppression would be the most effective strategy.
7.1 Introduction

The EU project xDELIA (Xcellence in Decision-making through Enhanced Learning in Immersive Applications, www.xdelia.org) aims to provide games for training to improve the decision quality of private investors, based upon the use of biofeedback to learn techniques of emotion regulation. The underlying hypothesis is that investors’ perform better when they can regulate their emotions with skill, thus be able to have less negative emotions. This is much like the sunny weather effect, which states that during days with good weather, people are in general more positive, and stock returns are stronger (Hirshleifer and Shumway, 2003).

7.2 Brain-Computer Interaction

Research in the field of Human-Computer Interaction (HCI) is addressing the development and evaluation of many less well established interaction technologies, from haptic, olfactory, and gestural interfaces (Kortum, 2008) to augmented reality (Linaza et al., 2012; Takacs et al., 2008) as well as interaction based on neuroscience, i.e. Brain-Computer Interfaces (BCI) (Minnery and Fine, 2009). BCI, as defined by (Pfurtscheller et al., 2010), has four criteria: (I) the recording device must rely on signals recorded from the brain; (II) there must be at least one recordable brain signal that the user can learn how to intentionally manipulate; (III) real-time processing must be available; and (IV) the user must obtain feedback. Note that this does not exclude systems that utilize additional input (e.g. mouse or keyboard) as well as recorded data from the brain. BCI has been used extensively to help disabled people (e.g. Guger et al., 2009; Hoffmann et al., 2008; Wolpaw et al., 2002) and some applications exist for games for a general audience. Maybe the most famous game that utilizes BCI is Brainball (Hjelm, 2003) where the two players each wear an electroencephalogram (EEG) that measures their relaxation. The game has a ball that is situated in between the two players, that moves towards the player that is the least relaxed. The winning condition is to send the ball across to the opposing player’s end of the play surface.

Since Brainball, many games and game technologies using BCI have been developed. BCI games may use one or both of explicit and implicit feedback (Kuikkaniemi et al., 2010), where implicit means that signals from the sensors modulate the game without the knowledge of the player. Explicit feedback means that the player can perceive the feedback, and use that information to modulate or control the gameplay. Kuikkaniemi et al. (2010) recommended the use of explicit
feedback since that increases immersion, positive affect, and players are able to manipulate their own physiological state. (Zander and Kothe, 2011) discussed the notion of active, reactive, and passive BCI, where active constitutes direct conscious control, without being affected by external stimuli. Reactive on the other hand is affected by external stimuli, but is indirectly modulated by the player in order to control the game. Passive derives its outputs from arbitrary brain activity without conscious control.

Nacke et al. (2011) implemented a game that used different forms of biophysical sensors, namely eye tracking, galvanic skin response, electromyography, electrocardiography, respiration rate, and temperature. The game they created used these devices as either passive or active inputs. They concluded that these types of interaction devices should be used to augment games in addition to more common interaction devices. Moreover, they state that active devices should preferably be mapped naturally to reflect an action in the virtual world, while passive devices are more suited to influence environment variables to enhance gameplay.

The work presented here used BCI, more specifically the Emotiv Epoc\(^1\), a 16-channel EEG device specifically developed in order to be a consumer-friendly device.

### 7.2.1 Effects of Emotion Regulation

Working memory resources consist of a visuo-spatial sketchpad, a phonological loop, and a central executive (Baddeley, 1992). The central executive controls attention, dictating what to inhibit and what to attend to. It also has limited resources, meaning that not everything can be attended to at once; attention has to be prioritized (Baddeley, 1996).

A task can be defined in multiple ways, namely: I. by the pattern of stimuli it emits; II. as the behaviour required in order to perform successfully on the task; III. as descriptions of the typical task behaviour that are exhibited during task performance; IV. as skills required to perform tasks (Wood, 1986). Throughout this article a task is described as II, the behaviour required in order to achieve a certain level of performance. (Beal et al., 2005) describe peoples’ effectiveness on tasks in terms of the amount of cognitive resource they can allocate to them. A completely effective person performing a single task would allocate 100

**Situation selection**, i.e. putting oneself in situations that are comfortable and non-upsetting, e.g. a person who is suffering from gephrophobia (fear of

\(^1\)www.emotiv.com
bridges) might travel great distances in order not to cross bridges on the path to their destination;

**Situational modification** , or problem-focused coping, involves modifying the situation in order to experience (or not experience) a specific emotion. E.g. in the previous example, if the person has to cross a bridge she might turn the volume of the radio up really high or start talking to her friend, thus altering how much attentional pull the environment has;

**Attentional deployment** to direct where the situation attention is placed. E.g. when the gephyrophobiac with a radio is walking over a bridge, she may try to remember the lyrics of the song being listened to, instead of attending to the fact that she is walking over a bridge;

**Cognitive change** is about selecting the meaning of the attended aspect of the situation, such as instead of thinking "this bridge is going to crumble down", thinking instead "there are very few people and cars on the bridge right now compared to usual circumstances, so I am certainly going to be fine";

**Response modulation** is the effort to affect emotions after they have already been elicited, e.g. trying to hide the stress emotions that occur when walking across a bridge.

The two categories relevant in this work are cognitive reappraisal (henceforth referred to just as reappraisal), which is a type of cognitive change defined as "construing a potentially emotion-eliciting situation in nonemotional terms", and expressive suppression (henceforth referred to just as suppression), which is a type of response modulation defined as "inhibiting ongoing emotion-expressive behavior" (Gross, 2002). Reappraisal generally reduces emotional experience and behaviour expression, and has no impact on memory. Suppression, on the other hand, generally reduces behaviour expression, but emotional experience is unaffected and memory is impaired (Gross, 1998a; Richards and Gross, 2000).

Wallace et al. (2009) used a game in order to assess if suppression or reappraisal traits were related to task performance in a digital game. They found that reappraisal tendencies correlated with better performance and suppression with worse performance. They explained their results by stating that the cognitive load was higher for suppression than from reappraisal, since reappraisal happens before the onset of an emotion, rendering the emotional response lower and thus creating less off-task attentional distraction. It has been shown in other experiments that reappraisal lowers responses such as disgust compared to suppression (Gross, 1998b).

Improved expertise development requires that a person is motivated, has a
good teacher (or teaching agent) and spends time on deliberate practice (Ericsson, 2006). The implemented game described in this paper is intended to provide a base for deliberate practice of emotion regulation in conjunction with a teacher. In order to assess the effectiveness of the use of this game in this way, the game needed to be tested to see if emotional regulation has any affect on game performance (i.e. game score and accuracy). This leads to the following hypothesis:

- Emotion regulation strategies or tendencies influence game performance.
  Based on Gross (Gross, 1998a; Wallace et al., 2009), the reappraisal strategy appears to be beneficial and is presumed to yield better results and more effective emotion regulation in this application than the suppression strategy. This leads to the further hypotheses:

- The emotion regulation strategy reappraisal will yield less time spent in unwanted arousal states compared to suppression.

- Reappraisal will yield better results, compared to suppression, in accuracy when arousal levels are the same due to more cognitive resources being available when using reappraisal.

The xDELIA-project aims at training emotion regulation strategies in order to encourage good financial behaviour. It is reasonable to expect that a game that gives instantaneous feedback on a player’s emotional state provides a good pedagogical approach since direct feedback has been shown to be good in task learning situations (Hattie and Timperley, 2007). In order to make this a priority within
the game, the players’ emotional state directly influences the difficulty of the game. (Lindley and Sennersten, 2008) hypothesize that a rewarding flow state in gameplay is associated with attentional demand, occurring when task-oriented schema execution demands attentional resources above a level that would result in player boredom and below a level that would result in excessive difficulty and consequent frustration. This suggests an optimal zone for player reward in gameplay, which is not necessarily (or typically) achieved by minimising game difficulty. Hence in the Aiming Game described here, modifying game difficulty based upon a player’s level of physiological arousal may allow players to optimise their gameplay experience by using emotion regulation to tune the difficulty level to create a level of attentional demand more likely to result in an experience of flow during play.

7.3 Method

7.3.1 Eeg

The Emotiv EPOC Neuroheadset (www.emotiv.com), a wireless electroencephalography (EEG) headset was used for collecting EEG data. The EPOC collects data from 14 saline sensors that rest against the scalp. This data is then processed to obtain signals representing affective, cognitive and expressive states, that are available for use in application software in real time through a standard developer kit (SDK). During this experiment, only the signal interpreted as the affective state of instantaneous excitement, or arousal, was used, with arousal represented by a value between 0.00 and 1.00.

7.3.2 The Game

The game used in this study, The Aiming Game (see fig 2)(Cederholm et al., 2011), is a two dimensional shooting game, with a fixed visual display resolution of 1280x1024 pixels, that consists of four levels, each providing 180 seconds of play time. The goal within these levels was to shoot down as many black airplanes as possible, which spawned consistently during the game at a rate of one airplane per 0.8 seconds and had speeds randomly assigned between 100-200 pixel/second. From the second level and onwards, distractors in the form of red airplanes, which the player should not shoot down, were also spawned, but with a rate of one airplane every 0.4 seconds with speeds randomly assigned between 130-260 pixel/second. Each black airplane that was shot down gives the player ten points;
red airplanes subtracted ten points from the player score, and each shot cost two points.

Game difficulty was modulated by player arousal level (acquired through the Emotiv EPOC), so players needed to regulate their emotions while playing the game in order to maximise their score. Difficulty was reduced with a low arousal value during the first three levels of the game and with a high arousal value during the fourth level. Arousal was scaled into five discrete categories, 1-5, based upon the value acquired from the EPOC. The further away from the desired arousal value the players were, the harder the game became. The game was made harder by randomly misplacing the aiming cursor, in a smooth motion, from the point controlled by the mouse, where further away from the target arousal level resulted in a bigger displacement, and also by blurring the airplanes (see fig 3.), where further away from the target arousal level resulted in increased blurring. The formula for the aiming offset was $30\times (\text{arousal} - 1) \times R$ pixels for both X- and Y-coordinates, where $R$ was a random number between either -1 and -0.5 or 0.5 and 1. This created a square in which the aiming cursor moved randomly, while at the same time ensuring that it would not move around in the middle of the square, due to the value of $R$. The aiming misplacement happened every 0.5 seconds (see fig 4.). In addition to the changing difficulty of the game features, players also had a bar on the left side of the screen that represented their arousal value. At all times the score of the players was visible in the bottom right corner.

### 7.3.3 Participants

Twenty-two subjects were recruited for a trial run and functioned later as a control group. Later, thirty-two more subjects were recruited and evenly divided into the two test groups, instructed in one each of the suppression and reappraisal emotion regulations strategies. Because of inbuilt technical characteristics of the EPOC, six of the subjects’ data could not be used for further analysis.

### 7.3.4 Questionnaire

The Emotion Regulation Questionnaire (ERQ) (Gross and John, 2003) was used in order to assess participants’ tendencies to use the two broad emotion regulation categories suppression and reappraisal. All questions were administered through a web tool (websurvey.textalk.se) where participants answered six questions about reappraisal tendencies and four questions about suppression tendencies. All these questions were answered on a seven point Likert scale, where 1 meant "strongly disagree", 4 meant "neutral", and 7 meant "strongly agree". In addition to the
7.3. METHOD

Figure 7.2: The second level of the Aiming Game, during minimum arousal

Figure 7.3: The second level of the Aiming Game, during maximum arousal
Figure 7.4: The outer bounds of where the cursor was misplaced to during maximum arousal
ERQ, demographic data was collected regarding age and gender. The factors "shooting game experience" and "arousal game experience" were also collected on a ten point scale where 1 was "very little" and 10 was "very much". After the experiment participants were asked to report on how they thought they handled their emotions as well as ranking the various game levels in terms of how difficult they were experienced to be.

7.3.5 Procedure

The participants were asked to fill in the web questionnaire at their own pace. After this the EPOC was applied, started, and tested. At this point and before starting the game, the suppression and reappraisal groups received their respective written instructions on how they were supposed to manage their emotions during the play session, while the control group just started the game. The game was played for 720 seconds and afterwards the participants from the test groups filled in a questionnaire containing questions about how well they performed the handling of emotions and ranking the different levels according to how difficult they were perceived to be. Additional comments were recorded for possible use in enhancing the game in the future as well as to explain possible odd results.

7.4 Results

The factors used in the analysis were Experience in shooting games; ERQ reappraisal score; ERQ suppression score; and condition (test, reappraisal, and suppression). The game statistics used were mean arousal for each phase, which was calculated by multiplying the time spent in each arousal state by the state number 1-5 and then dividing the time by 180 (the play time of a level). Accuracy in gameplay was calculated for each arousal level and over three phases (the forth one being handled separately since the arousal was reversed) by taking the total number of aircraft hits divided by the total number of shots fired. Game score was also considered when making the analysis.

An analysis of variance (ANOVA) was performed with regards to mean arousal time for each phase, but no significant differences were found between the groups (N = control 21, suppress 15, reappraise 12). Since there was a difference in the experience of shooting games between the groups (p= 0.001, ANOVA one-way) it is hard to draw any conclusions regarding differences in accuracy and game points arising from the use of different emotion regulation strategies by the groups. Participants reported that it was generally hard to uphold the strategy
they were instructed to use (mean 3.22 on a 7-point Likert scale) and six subjects reported that they had abandoned the strategy after a while when they focused on the game.

Despite the lack of correlations (Pearson’s correlation was used throughout this work) between game score and the manipulations in the experiment, a correlational analysis of all subjects (i.e., all groups combined) revealed strong correlations between experience with shooting games and accuracy as well as suppression and accuracy (see Table 1).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.182</td>
<td>.354**</td>
<td>.561**</td>
<td>.647**</td>
<td>.614**</td>
<td>.645**</td>
<td>.475**</td>
</tr>
<tr>
<td>2. Reappraise</td>
<td>1</td>
<td>.227</td>
<td>.033</td>
<td>.158</td>
<td>.183</td>
<td>.085</td>
<td>.289*</td>
</tr>
<tr>
<td>3. Suppress</td>
<td>1</td>
<td>.361*</td>
<td>.459**</td>
<td>.483**</td>
<td>.434**</td>
<td>.469**</td>
<td>.469**</td>
</tr>
<tr>
<td>4. Accuracy Arousal 1</td>
<td>1</td>
<td>.759**</td>
<td>.658**</td>
<td>.564**</td>
<td>.524**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Accuracy Arousal 2</td>
<td>1</td>
<td>.942**</td>
<td>.780**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Accuracy Arousal 3</td>
<td>1</td>
<td>.813**</td>
<td>.573**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Accuracy Arousal 4</td>
<td>1</td>
<td>.415**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Accuracy Arousal 5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Correlation is significant at the 0.05 level (2-tailed).
**Correlation is significant at the 0.01 level (2-tailed).

Table 7.1: Pearson’s correlations between experience, ERQ results, and accuracy for all participants (N=48).

Since there was a correlation between suppression and experience with shooting games, new groups based on these two variables were formed. First, a division was made between participants with suppression less than 4 (N=22) and participants with suppression 4 or higher (N=26). In these groups there were still correlations between shooting games and accuracy, but none between suppression and accuracy. A second division was made by experience of shooting games; the groups were beginners (experience 5 or less) and experts (experience 6 or more), thus making two groups of 24 subjects in each category. T-tests between these groups showed no statistically significant difference in either of the
7.5. DISCUSSION

ERQ results or mean arousal during any phase, making the two groups easily comparable. Correlation data for beginners and experts for shooting game experience, reappraisal, and suppression, as well as accuracy can be seen in table 2. T-tests between experts and beginners showed significant differences between accuracy during all different arousal states, and between game scores during all phases, but no statistically significant differences in arousal data from the EPOC (see table 3).

<table>
<thead>
<tr>
<th></th>
<th>Beginner</th>
<th>Expert</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exp SG</td>
<td>Reappr</td>
</tr>
<tr>
<td>Exp SG</td>
<td>1</td>
<td>.256</td>
</tr>
<tr>
<td>Reappr</td>
<td>.256**</td>
<td>1</td>
</tr>
<tr>
<td>Suppr</td>
<td>.621</td>
<td>.144</td>
</tr>
<tr>
<td>Acc 1</td>
<td>.382</td>
<td>.157</td>
</tr>
<tr>
<td>Acc 2</td>
<td>.568**</td>
<td>.265</td>
</tr>
<tr>
<td>Acc 3</td>
<td>.594**</td>
<td>.363</td>
</tr>
<tr>
<td>Acc 4</td>
<td>.632**</td>
<td>.410*</td>
</tr>
<tr>
<td>Acc 5</td>
<td>.553**</td>
<td>.444*</td>
</tr>
<tr>
<td>Inv Acc 1</td>
<td>.465</td>
<td>.616*</td>
</tr>
<tr>
<td>Inv Acc 2</td>
<td>.057</td>
<td>.386</td>
</tr>
<tr>
<td>Inv Acc 3</td>
<td>.413</td>
<td>.635*</td>
</tr>
<tr>
<td>Inv Acc 4</td>
<td>.164</td>
<td>.215</td>
</tr>
<tr>
<td>Inv Acc 5</td>
<td>.688**</td>
<td>.596**</td>
</tr>
<tr>
<td>Ph. 1 Scr</td>
<td>.485*</td>
<td>.314</td>
</tr>
<tr>
<td>Ph. 2 Scr</td>
<td>.558**</td>
<td>.215</td>
</tr>
<tr>
<td>Ph. 3 Scr</td>
<td>.393</td>
<td>.308</td>
</tr>
<tr>
<td>Ph. 4 Scr</td>
<td>.262</td>
<td>.518*</td>
</tr>
<tr>
<td>Tot Score</td>
<td>.497**</td>
<td>.303</td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.05 level (2-tailed).
** Correlation is significant at the 0.01 level (2-tailed).

Table 7.2: Pearson’s correlational data between shooting game experience, ERQ score, inversed arousal accuracy, and game score for each phase.

7.5 Discussion

The first hypothesis that Emotion regulation strategies or tendencies influence game performance is confirmed, since the ERQ score of suppression correlated strongly with accuracy (see table 1). When scrutinizing the results further it was shown that the correlations existed in the expert group, but stronger than
### Table 7.3: Mean data for experts and beginners, as well as t-test results between the groups.

<table>
<thead>
<tr>
<th></th>
<th>Sig. (2-tailed)</th>
<th>Mean Experts</th>
<th>Mean Beginners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acc 1</td>
<td>.000</td>
<td>.5678</td>
<td>.3486</td>
</tr>
<tr>
<td>Acc 2</td>
<td>.000</td>
<td>.4960</td>
<td>.2779</td>
</tr>
<tr>
<td>Acc 3</td>
<td>.000</td>
<td>.4114</td>
<td>.2382</td>
</tr>
<tr>
<td>Acc 4</td>
<td>.000</td>
<td>.3913</td>
<td>.1967</td>
</tr>
<tr>
<td>Acc 5</td>
<td>.009</td>
<td>.3243</td>
<td>.1832</td>
</tr>
<tr>
<td>Acc Inversed 1</td>
<td>.031</td>
<td>.4354</td>
<td>.1909</td>
</tr>
<tr>
<td>Acc Inversed 2</td>
<td>.011</td>
<td>.2748</td>
<td>.1579</td>
</tr>
<tr>
<td>Acc Inversed 3</td>
<td>.004</td>
<td>.3396</td>
<td>.1934</td>
</tr>
<tr>
<td>Acc Inversed 4</td>
<td>.004</td>
<td>.3599</td>
<td>.2095</td>
</tr>
<tr>
<td>Acc Inversed 5</td>
<td>.037</td>
<td>.4852</td>
<td>.2950</td>
</tr>
<tr>
<td>Ph. 1 Total Scr</td>
<td>.000</td>
<td>224.0000</td>
<td>-2.9167</td>
</tr>
<tr>
<td>Ph. 2 Total Scr</td>
<td>.000</td>
<td>249.0833</td>
<td>-4.8333</td>
</tr>
<tr>
<td>Ph. 3 Total Scr</td>
<td>.000</td>
<td>302.5833</td>
<td>12.0000</td>
</tr>
<tr>
<td>Ph. 4 Total Scr</td>
<td>.004</td>
<td>103.4737</td>
<td>-163.3000</td>
</tr>
<tr>
<td>Total Score</td>
<td>.000</td>
<td>775.6667</td>
<td>4.2500</td>
</tr>
<tr>
<td>Ph. 1 Weight Arousal</td>
<td>.436</td>
<td>2.7694</td>
<td>2.6875</td>
</tr>
<tr>
<td>Ph. 2 Weight Arousal</td>
<td>.679</td>
<td>2.9347</td>
<td>2.9972</td>
</tr>
<tr>
<td>Ph. 3 Weight Arousal</td>
<td>.537</td>
<td>2.8039</td>
<td>2.9238</td>
</tr>
<tr>
<td>Ph. 4 Weight Arousal</td>
<td>.959</td>
<td>3.0532</td>
<td>3.0661</td>
</tr>
</tbody>
</table>

Table 7.3: Mean data for experts and beginners, as well as t-test results between the groups.
7.5. DISCUSSION

initially shown (see table 2). Experience with shooting games was also correlated with game performance, as might be expected.

The further hypotheses, that:

- The emotion regulation strategy reappraisal will yield less time spent in unwanted arousal states compared to suppression.
- Reappraisal will yield better results, compared to suppression, in accuracy when arousal levels are the same due to more cognitive resources being available when using reappraisal.

were disconfirmed, since reappraisal did not result in less time spent at lower arousal levels, nor a conclusively better score or accuracy while on the same arousal levels as suppressors.

The difference between experts and beginners could be explained by the two groups in fact performing two different cognitive tasks: the beginners are struggling to get the crosshair to the right position, while the experts are used to this task and are instead focusing on shooting down aircraft and/or managing their emotions in order to perform better. This implies that a certain level of proficiency is required in shooting in order to be able to train emotion regulation with this game.

Suppression is only correlated with accuracy in the expert group while the (non-inverted) arousal values were between 2-4. This can be explained by suppression not being needed when the arousal level is 1 (since no distortion or aiming offset occurs) because there are no unwanted emotions to regulate. At the other extreme, at arousal level 5, there might be too much arousal to be able to regulate, rendering the suppression strategy irrelevant for performance.

For experts, when the impact of the arousal value was inverted (i.e. a high arousal value was beneficial) the only correlation that could be seen was with suppression and the middle arousal value, where $r=-.493$. This seems like a randomly misplaced value since all other correlations (although non-significant) with suppression are positive.

The relationship between experience with shooting games, game score and accuracy among beginners seems very clear. The more experienced the player the better the performance in the non-inverted arousal case in relation to accuracy, except at the lowest arousal level. This seems natural, since it requires more skill to aim with a randomly changing aiming offset than without any offset. In the inverted case there is no relationship between these variables except in the highest arousal condition, which can be explained by better players seizing the opportunity to catch points late in the game. The total game score correlated
strongly with experience, but not the third and fourth phases of play. This might have to do with learning effects, where experience might only have played a role early in the game, while later in the game all players adapted and performed at better levels.

A possible reason why the tendency to suppress was superior to the tendency to reappraise within the expert group might be that the game situation was fairly simple. In the step where reappraisal (cognitive change) can occur, there are actually very few options to chose from since the game environment will tell the player what is important by having a scoring mechanism. Compared to everyday life situations, which include things like complex social interactions, this is a very simple situation in which there might be less room for positive interpretation. Suppression is always focused upon the emotions themselves that a subject is experiencing, making them independent of whether a situation is simple or complex.

The results from this experiment suggest that for experienced players, the most relevant emotion regulation strategy is suppression. As a training device this means that the skills that will be trained are most likely the shooting ability and thereafter the ability to suppress effectively. Since this is not what the investors strive towards, it is probably not the most effective training platform for them. However, there are cases, such as when the hiding of emotions is necessary, when good suppression is relevant in which this game could be an excellent training tool. Moreover, reappraisal is about reformulating a situation into a non-emotional one, which requires knowledge about how to reformulate it. This could be theorized to be harder when approaching novel situations. Suppression is a reaction to the emotional response and thus much more general. Hence it can be argued that in novel situations, and during short periods, suppression might be superior to reappraisal. Also, since there are no good tools for training reappraisal, suppression can be a good substitute, if encouraged to not be used consistently during a day because of adverse long-term effects (Gross, 1998a), since training a skill generally makes people better at it, thus requiring less cognitive effort.

For HCI and game design, this has some tangible consequences. Generally, computer systems should be designed to avoid negative emotions. When they fail to do so, they need to help people manage their emotions in a way that is beneficial for the user’s long term health.

Game design (and possibly other software as well) needs to take into consideration what emotions regulation strategies people are using and being taught (if any) through playing games. In the case of the Aiming Game here, it is not a significant issue since the players were briefed about emotional effects and did
not play for a long duration. However, several games can be, and are, played for many hours. Without knowing the consequences and learning of emotion regulation strategies, this may have a very adverse effect on people’s lives. It is also an option that games are teaching people to regulate their emotions by release through virtual (possibly violent) behaviour. More research is needed in this regard to conclude whether this is applicable in non-Bci games or not.

7.6 Conclusion

Although no difference was found between the beginner and expert groups in terms of arousal states, very different correlational patterns emerged when emotion regulation and accuracy were brought into the picture. The game may be suitable for training the expressive suppression strategy, but not reappraisal.
CHAPTER 7.  A STUDY OF THE EFFECT OF EMOTION REGULATION IN A SHOOTING GAME USING BIO-FEEDBACK FROM ELECTROENCEPHALOGRAPHY FOR TRAINING IN EMOTION REGULATION
Chapter 8

Integrating Biosignals into Information Systems: A NeuroIs Tool for Improving Emotion Regulation

Abstract

Traders and private investors are aware that emotional processes can have material consequences on their financial decision performance. However, typical learning approaches for de-biasing fail to overcome emotionally driven financial dispositions; mostly because of subjects’ limited capacity for self-monitoring. Our research aims at improving decision makers’ performance by (i) boosting their awareness to their emotional state and (ii) improving their skills for effective emotion regulation. To that end we design, implement and evaluate a serious game based NeuroIs tool that continuously displays the player’s individual emotional state, via biofeedback, and adapts the difficulty of the decision environment to this emotional state. The design artifact is then evaluated in two laboratory experiments. Taken together, our study demonstrates how Is design science research can contribute to improving financial decision making by integrating tools from cognitive neuroscience in IT artifacts.
Chapter will be available at later point, due to Journal embargo restrictions.
Chapter 9

Identifying the Moment when Demanding Cognitive Load Overshadows the Arousal Effect on Pupil Dilation During a Stressful Task

Abstract

There is a need to develop applications tailored for individual needs and skills. Such applications would need to know when the individual user would perform optimally and when would he be overwhelmed, such that the task could be tailored individually for that specific user. Our research explores the moment when demanding cognitive load overshadows the arousal effect on pupil dilation during a stressful task. To identify this exact moment, PD and electrocardiogram (ECG) data were collected from participants playing a serious game as a dynamic learning environment displaying biofeedback based on ECG measurements; further-
more, the Pd data was analyzed in relation to the heart rate (HR) data during this stressful task. Peak of sympathetic nervous system activity was identified in the Pd data; thus, this data coupling enabled observation of sympathetic and parasympathetic pathway activation and dominance over each other.
9.1 Introduction

This research was carried out as part of the international project xDelia (Xcellence in Decision-making through Enhanced Learning in Immersive Applications) funded from the European Commission under the 7th Framework Programme. Under the xDelia project we developed a serious game as a dynamic learning environment displaying biofeedback based on psychophysiological measurements that aim at improving the player’s emotional awareness and training of his/her ability to down-regulate levels of high arousal. When we talk about the arousal, we are referring to an emotional arousal measured using psychophysiology, throughout this paper referred to as psychophysiological arousal. An advantage of such a tool is that it measures reliable emotional arousal in a stressful environment (Schaaff et al., 2011; Jerčić et al., 2012; Astor et al., 2014) and the result data is not biased by interoception, as is the case of questionnaire assessment (Elmes et al., 2011; Schwerdtfeger, 2004). In the game, the player faces a financial decision scenario that is directly linked to his or her individual emotional state which is also indicated via biofeedback. Based on heart rate measurements, the game incorporates the player’s arousal level into the game task making it more (or less) difficult as the game continuously adjusts the financial decision scenario (Astor et al., 2014). Depending on the player’s ability to regulate arousal emotions, s/he could perform better (or worse) in the game task, which is on its own demanding enough. The player wears an electrocardiogram (ECG) sensor that transfers the data via Bluetooth to the game. This technique requires physical contact of one or more sensors with the body of a user. For broader commercial acceptability, non-contact methods for measuring indicators of user arousal are highly desirable, since these can be used with reduced setup time, no added inconvenience to users and without tight constraints upon motion during human-computer interaction. The main interest of such a method is that it is maximally non-contact, which allows a relatively ecological situation. Following those propositions an infrared eye-tracking system and a mobile wireless ECG sensor, emerged as a solution providing quantitative measurements of activation underlying central nervous system mechanisms (parasympathetic and sympathetic pathway) in order to pinpoint the moment when demanding cognitive load overshadows the arousal effect.

The pupil is an opening at the center of the iris surrounded by the iris muscles that respond to outside stimuli to control the amount of light that reaches the retina. We differentiate between two types of iris muscles: the sphincter muscle sphincter pupillae forming a band around the inner margin of the pupil and radial muscle dilator pupillae. Two synergistic pathways control pupil size and dynamics. Those pathways are the parasympathetic pathway and the sympathetic pathway, operating on the smooth muscles of the pupil. The parasympathetic pathway
CHAPTER 9. IDENTIFYING THE MOMENT WHEN DEMANDING COGNITIVE LOAD OVERTURNS THE AROUSAL EFFECT ON PUPIL DILATATION DURING A STRESSFUL TASK

is mediated with the efferent pathway originating in the Edinger-Westphal oculo-motor complex in the midbrain and innervates the sphincter which is the circular muscle responsible for constriction, as seen in the reflex reaction to light. Inhibition of Edinger-Westphal complex results in relaxation of the sphincter muscles and, thus, significant dilation. The sympathetic pathway, mediated by the hypothalamus, innervates the radial dilator muscle of the iris responsible for dilation (Steinhauer et al., 2004; Privitera et al., 2008). The described mechanism has been illustrated on the schematic view on Figure 1.

![Figure 9.1: Model of the Ans control pathways for the pupil and the heart](image)

The concept of physiological arousal has been generally validated in models of emotion (Cannon, 1994; Porges, 1995; Morrison, 2001). Pupil dilation is influenced by physiological arousal, together with attention and interest (Andreassi, 2007). It was therefore hypothesized for our study reported here that pupil dilation could provide a useful measure of physiological arousal (Granholm and Steinhauer, 2004). Early explorers (Hess and Polt, 1964) viewed the pupillary response as reflecting level of arousal or emotionality; followed a few year later
(Hess and Polt, 1964; Kahneman and Beatty, 1966) by studies arguing that the pupillary response might reflect cognitive activity as well as or instead of arousal. Thus Stanners et al. (1979) concludes that the pupil response will show an arousal effect only when the cognitive demands of the situation are minimal. If the situation requires a substantial level of cognitive activity, only this will be indicated by the pupillary response overshadowing the emotional arousal effects. Stein- hauer et al. (2004) gave neuroscientific support for this theory where they report that multiple pathways impinge on the Edinger-Westphal complex, resulting in pupillary dilation through inhibition of the parasympathetic pathway. Those multiple pathways that generally increase in inhibition across all task conditions may well include contributions of both direct cortical/indirect cortico-thalamic- hypothalamic pathways (Lowenstein, 1955) and reticular pathways contributing to arousal (Bonvallet and Zbrozyna, 1963). It has been shown that the pupil size reflects processing load or mental effort (Moresi et al., 2008; Recarte and Nunes, 2003). Thus Steinhauer et al. (2004) states that demanding cognitive load, most likely associated with frontal cortical functioning, contributes heavily to this inhibitory process overshadowing the arousal effect. Such proposition had prompted exploration for the moment when the overshadowing occurs.

Granholm and Steinhauer (2004) state that changes in pupillary motility have been noted and employed as indicators of both emotional arousal and signs of medical state for over two millennia of history. A proposition came from Partala and Surakka (2003) to use pupil size variation as a computer input signal in affective computing, for example. Bradley et al. (2008) present a strong case supporting that the pupil's response during affective picture viewing reflects emotional arousal associated with increased sympathetic activity. They argue that emotional arousal is a key element in modulating the pupil's response. Furthermore, they state that rather than varying with cardiac deceleration mediated by differences in parasympathetic activity, they found that pupillary changes covaried with skin conductance reactions, providing collateral support for the hypothesis that pupil diameter during emotionally arousing picture viewing predominantly reflects sympathetic nervous system activity. While in the studies reported in the previous paragraph that have explored effects of mental load and sustained cognitive processing on pupil change, the observed pupillary dilation appeared to be mediated by parasympathetic inhibition of the sphincter muscle. These findings support the hypothesis that pupillary changes during affective picture viewing are mediated by increased sympathetic activity and strongly suggest that emotional arousal affects the pupil dilation, independent of whether pictures are pleasant or unpleasant in hedonic valence. These insights support the Steinhauer et al. (2004) suggestion that it is unclear to which extent is the cognitive activity uniquely associated with sympathetic or parasympathetic nervous system activity. They further argue that, if we are to provide quantitative measurements
of activation underlying central nervous system mechanisms, it is important to differentiate between activation of these pathways. With that knowledge one can evaluate the neurophysiological systems that contribute to cognitive activities by monitoring pupillary dynamics.

Berntson et al. (2007) states that the heart rate is a measure of activity of both the sympathetic and parasympathetic autonomous nervous system. Correspondingly, sympathetic activity tends to increase heart rate, while parasympathetic activity decreases it. Previous studies have shown that skin conductance changes are greater when viewing pleasant and unpleasant, compared to neutral pictures. This suggests that this sympathetically mediated response covaries with emotional arousal (Lang et al., 1993). In contrast, cardiac deceleration is generally greater when viewing unpleasant, compared to either pleasant or neutral pictures (Bradley et al., 2001). This suggests that this is a parasympathetic activity (Berntson et al., 1989) as shown in the pharmacological blockade studies of fear bradycardia in animals. The term bradycardia defined as a heart rate which falls below 60 beats per minute. Similarly to the pupil, Berntson et al. (2007) further report that the extent interplay between sympathetic or parasympathetic nervous system activity provides a measure of a subject’s emotion regulation capabilities. Moreover, just like in the effects observed in the pupil, an increased task difficulty (e.g. difficulty of a mental arithmetic task) increases the heart rate (Boutcher and Boutcher, 2006; Sosnowski et al., 2004). So the same suggestion made by Steinhauer et al. (2004) is applicable to the heart, where it is unclear to which extent is the cognitive activity uniquely associated with sympathetic or parasympathetic nervous system activity.

According to the research reported by Yerkes and Dodson (1908), there is an optimal level of physiological arousal increasing with the task difficulty, which results in one’s maximum performance for a given task. Since \( P_d \) reflects subject’s activation and difficulty of a task for a certain individual, we could potentially identify the point of optimal arousal in the \( P_d \) data. The Yerkes-Dodson law established that performance increases with arousal up to a point at the peak of optimal arousal level, after which performance decreases as arousal continues to increase. As pupil dilates with the cognitive demand of the task, it constricts and flattens out when the cognitive demands for the task become overwhelming. At that point \( P_d \) should stagnate and start decreasing even though we should be able to observe increase of the arousal on the heart rate.

We explored activation of sympathetic and parasympathetic pathways to investigate when one prevails over the other. The \( P_d \) data was analyzed in relation to the heart rate (\( H_r \)) data during this stressful task. Peak of sympathetic nervous system activity was identified in the in the \( P_d \) data; thus, this data coupling enabled observation of sympathetic and parasympathetic pathway activation and
dominance over each other. That is, when cognitive demand overshadows the arousal effect on pupil dilation. The study used one of the games developed in xDelia, the Auction Game (Jerčić et al., 2012), where the game was interconnected with the eyetracker and ECG sensor. This interconnection enabled monitoring and logging of real-time pupil responses from the player in order to pinpoint the moment when demanding cognitive load overshadowing the arousal effect. It could later be used to provide feedback to the player to shape awareness of their own optimal arousal reaction.

9.2 Material and methods

Other work within xDelia addressed the use and validation of heart rate and heart rate variance as measures of physiological arousal. Since the current study was based upon the proposal to substitute pupil dilation for heart rate, a correlation study was conducted where a number of participants interacted with the same game task stimulus while continuous measures of heart rate and pupil dilation were made. Subsequent analysis of correlations between heart rate and pupil dilation would provide evidence for or against the hypothesis that pupil dilation could substitute for heart rate, indirectly providing evidence that it could provide a useful measure of physiological arousal. Definite correlation would support the hypothesis, in so far as heart rate can be correlated with arousal (Frankenhaeuser and Johansson, 1976).

9.2.1 Participants

Twenty-one students from Blekinge Technical Institute participated in the experiment. They all ranged an age between 20-24 years. Participants haven’t reported any ophthalmologic problems (other than corrected vision); furthermore, no psychiatric or major medical disorders were reported. Participants were paid for participating and were told that they could earn additional profit according to their actions during the experiment. All participants received complete information on the study’s goal, experimental conditions and gave their informed consent approved by the Local (BTH) Ethical Committee and EU ethics advisory board in the frame of FP7 and Cordis funding.
9.2.2 Materials

The Auction Game (Jerčić et al., 2012) was selected to function as the game stimulus for the study. The Auction Game was designed to support training in emotion regulation, but also to raise the player’s awareness of their emotions during play. In other words, it would also support the player in achieving mindfulness of emotions. To achieve this, in the default version of the game heart rate is used as a measure of arousal.

The objective of the Auction Game is to train the players in emotion regulation. Better emotion regulation leads to better awareness so one could regulate himself in future gameplay, thereby making it easier to achieve higher profits in the game. Achieving higher profits was supported by providing estimations given to the player from virtual in-game advisors. The player was presented with three price estimations before s/he had to reach a buy or sell decision (Figure 2). Price estimations were directly linked to the player’s emotional arousal level: they deviated from the correct price with higher variance the more aroused the player was. In the Auction game the gameplay task is to calculate a mean value of a stock, from the given estimations, to be able to sell or buy the stocks at the correct best price; thus a lower emotional arousal would make the stock price estimations closer to the correct price so that a buy or sell decision could be made with a more predictable outcome. Moreover, goal of play was to reach the highest level possible. Hence, the ending time of single levels and the complete game could vary, depending on player’s skills. Player’s skills to earn money in the game were strongly related to player’s skills to regulate his/her overall emotional arousal state. Good emotion regulation would help the player to be successful in the game. From level to level the tasks requiring regulation of emotion became more and more difficult as new elements were included in a piecewise manner. Among the new elements were additional information, time, auditory and visual constraints. A certain randomness of these elements also strengthened the immersion and excitement of the game.

Stimulus (Figure 3) was presented in this fashion, before the onset of cloud stimulus participant had a one second pause. Every cloud was presented on the screen for one second or a slightly shorter if the person was highly aroused. The participant then had three seconds to reach a decision (Buy, Sell or No Trade) after which the winnings feedback was presented in the first half of the pre-cloud resting period. The player traded a fictional good and only its price was depicted as relevant for the game. For each traded good the player got three price signals indicated on the screen. The mean value of these price signals was the correct underlying price in the next period of play. After the indication of these price signals the player had to come to a buy or sell decision for an offered price.
Depending on the player’s decision she could earn or lose money (on average 2.50 €) depending upon whether she is correct or not. The decision had to be made within a limited amount of time (3 seconds), otherwise the player was punished by a loss of money (-5 €). If the player’s assets fell under a certain barrier below the starting level, s/he was ’kicked out of the game’. With every decision the player had the option to pause and not to trade a stock this round. This was achieved by clicking the ”No Trade” button. This decision would not reward the player with any profit for that round, but the loss of money (-5 €) for not taking a decision could be evaded. Moreover, the player had a chance to relax and regulate their emotional arousal before progressing further with the game.

To advance from one level to the next, the player had to gain a certain amount of money to reach the current level’s profit goal. As soon as s/he hit this amount, a button popped up and s/he could proceed to the next level. While in the first levels the constraint for physiological arousal was to down-regulate the player’s emotion, later levels introduced more emotionally distracting elements to the game, making it more difficult to keep arousal levels low. Among these was highly arousing music and distracting pictures. High levels of arousal led to increased variance on price signals, moving price signals and fake elements. This was to worsen the player’s economic performance and hence their performance in the game.
CHAPTER 9. IDENTIFYING THE MOMENT WHEN DEMANDING COGNITIVE LOAD OVERSHADOWS THE AROUSAL EFFECT ON PUPIL DILATION DURING A STRESSFUL TASK

Figure 9.3: Layout of the Auction Game screen
9.2. MATERIAL AND METHODS

In order to determine the baseline for arousal, Hr was measured in a resting phase prior to the game where a participant was presented with a blue screen with a focus cross in the middle. For the pupil, a participant was presented with three different color still images with a 60 seconds individual exposure time from which a baseline has been collected.

9.2.3 Apparatus

Subjects were seated in a normal lit room with no direct light source. Pupil dilation was measured using a Tobii T60 Eyetracker. The T60 has a data acquisition frequency of 60Hz which means that tracked information is updated every 20 ms. The eyetracker has an x- and y- angular accuracy of 0.5°, while the maximum gaze angle can be up to 35 degrees from the line to the centre of the screen. The eyetracker is supplied with its own analysis software, Tobii Studio™ that can be used for basic analysis, while we used and analyzed the exported logged text data. The eye tracker can also be interfaced to bespoke software for analysis or gaze-contingent- or eye-controlled applications.

For computation of heart rate (Hr), the ekgMove sensor developed by Mo- visens was used, which records electrocardiographic (ECG) signals with high accuracy. In contrast to most other commercially available ECG devices, the sensor was attached to the chest using a flexible belt with dry electrodes. Therefore, it was less obtrusive than other devices and offers higher wearing comfort. The ECG signal was transmitted via Bluetooth to the xAffect software environment (Schaaff et al., 2011). The xAffect software offers a modular framework that allows processing of data from various input devices and transfer of derived values via TCP/IP to other applications like the Auction Game. To get information about the current arousal level of a person, the heart rate was computed within xAffect from the raw ECG signal. An algorithm to derive the current arousal level from heart rate information was also implemented in the xAffect framework. The arousal levels inferred from the online ECG signal were computed in relation to an ECG signal recorded in the baseline period, which was recorded before the game starts. An advantage of this system was that it measures emotional arousal reliably in stressful environments and is not biased by interoception.

The implementation purpose of the integrated game engine, eyetracker and ECG sensor was to be able to validate one data set in relation to the other. The data sets were not perfectly synchronized due to a lag introduced by the ekgMove sensor and because of natural physiological delays in eye response followed by heart response. The eye muscle controls the retina to attend to stimuli and the stimuli are visually perceived first, followed by the heart rate response. The
CHAPTER 9. IDENTIFYING THE MOMENT WHEN DEMANDING COGNITIVE LOAD OVERSHADOWS THE AROUSAL EFFECT ON PUPIL DILATION DURING A STRESSFUL TASK

Unity © pro game engine was used to implement the game and the Tobii T60 eyetracker and the ekgMove (Movisens ©) HR sensor were interconnected with the Unity pro game engine. The ekgMove ekgMove sensor was integrated via the xAffect software. The software Tobii Studio™ was used to run the experiment, so still images and the game were presented and sequenced automatically via commands from this software.

9.2.4 Procedure

Upon arrival, subjects were given general information about the experiment and a description explaining the Auction Game. Their written consent explaining psychophysiological measurements and providing anonymity, was obtained. Each participant was seated in a recliner in a small, sound-attenuated, natural lit room with no direct light and the chest band holding psychophysiological ECG sensor was attached. Participants were seated in front of and facing the eyetracker at a 50-60 cm distance from the screen. Prior to the experimental session a calibration of the eyetracker was performed where participants should attentively focus on 4 individual dots distributed on the screen, after which the participant was instructed to start with the experiment. Luminance of the screen and the light-conditions around the screen were measured. First the participant was presented with three still images for 3 minutes (60 seconds each). The still images were: a) Plain white, b) Plain grey and c) Plain black (Figure 4). We recorded a baseline with individual pupil size (BPS) variation for still images. Before starting the game task presentation a practice run tutorial of the Auction Game was presented in order to acquaint subjects with the task. Thereafter, subjects were presented with a blue screen (Figure 4d) and directed to relax for 5 min while the HR baseline measurement was acquired. Once these baselines were logged, the actual game task started. At the end of each level, subjects were asked to assign arousal ratings they had experienced during the level played using the Self-assessment Mannequin (SAM) whose values ranged from 1 – 5. Following the experiment, the sensor was removed and the participant was asked to fill out a post experimental questionnaire. The experimenter subsequently debriefed, paid, and thanked the participant.

9.2.5 Data analysis

Data exported off-line from the the Tobii Studio™ for individual trials, contained the values for the pupil diameter in millimeters. The software itself corrected for the short blink periods, but the data was further checked for the longer blink periods. A linear interpolation was then applied. Baseline diameter was defined
9.3. RESULTS

Figure 9.4: Color still images for pupil dilation baseline

as the average diameter during the 60s still image presentation. Baseline was first examined separately, after which it was used to normalize the pupil diameter data.

Data exported off-line from the xAffect software for individual trials, contained the heart rate values in beats per minute (bpm). The xAffect Software performed a digital trigger detecting R-waves applied to the ECG signal to obtain interbeat intervals which were then reduced to heart rate. The software itself executed filtering and correction of data.

There was a discrepancy in the pupil diameter and the heart rate data sampling with 60Hz for the eyetracker and 1Hz for the xAffect software. This prompted creation of two pupil/heart datasets. One dataset had the heart rate upsampled to match the sampling rate of the eyetracker using linear interpolation. The other dataset had pupil diameter data downsampled to an average of 60 data points to match the sampling rate of the xAffect software. For both of those data sets first and second derivatives from the pupil diameter and heart rate data were calculated using the Matlab software; furthermore, the initial, first and the second derivative data were then analyzed for the correlation using the Pearson’s correlation coefficient in the R statistical software package.

In the experiment, data was analyzed as a repeated measures design for the factors of game task condition.

9.3 Results

We randomly allocated 21 participant aged between 20-24 years for the study using a serious game as a dynamic learning environment displaying biofeedback based on psychophysiological measurements that aims at improving the player’s emotional awareness and training of his/her ability to down-regulate levels of high emotional arousal. Goal of the study was to explore substituting heart rate psychophysiological arousal for pupil dilation measurement for the game task
presented (Figure 2).

### 9.3.1 Baseline values for pupil dilation

We calculated mean values for both the left and the right pupil dilation during color still images exposure. Absolute values in millimeters are presented in Table 1. One participant had an extreme loss of data during the baseline reading and he was excluded from this evaluation. The total minimum and total maximum values for each pupil show the range the diameter has among the set of participants when exposed to the baseline stimuli. These results are used to form a normalized scale for subsequent pupil dilation analyses. We averaged pupil dilation maxima, minima and mean absolute values for each eye averaged across all participants for the baseline measurements for each pupil and each participant for the white, grey, black and blue screens, indicating where the center values are. The values presented below in the Table 1 include the median and mean values for each pupil and each participant for the white, grey, black and blue screens, indicating where the center value is. One can notice that the maximum change to mean pupil diameter due to changing the screen color and brightness in the baseline study is $D_{\text{max}} - D_{\text{min}} = 3.84 \text{ mm} - 3.02 \text{ mm} = 0.82 \text{ mm}$. Note that the maximum pupil dilations measured were $BL_{\text{min}} = 1.34 \text{ mm}$; $BL_{\text{max}} = 6.15 \text{ mm}$.

<table>
<thead>
<tr>
<th>Stimulus screen</th>
<th>Pupil median/mm</th>
<th>Pupil mean/mm</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>3.02</td>
<td>3.025</td>
</tr>
<tr>
<td>Grey</td>
<td>3.27</td>
<td>3.28</td>
</tr>
<tr>
<td>Black</td>
<td>3.82</td>
<td>3.785</td>
</tr>
<tr>
<td>Blue</td>
<td>3.84</td>
<td>3.84</td>
</tr>
</tbody>
</table>

Table 9.1: Pupil median and means for both eyes together

### 9.4 Parasympathetic vs. sympathetic pathway activation behavior

Furthermore, we explored how the Pd behaves in relation to the Hr data. More specifically, the question investigated was: can we observe the pattern of behavior of Pd for a certain value of the Hr. Moreover, can we infer the moment of activation and dominance of parasympathetic over sympathetic pathway, as proposed by Steinhauer et al. (2004), where demanding cognitive load overshadows the arousal effect. To explore this we used the SPSS statistical
9.5 DISCUSSION

analysis package and grouped the absolute and the normalized Pd responses according to the accompanying HR data measured. The absolute and normalized Pd values were then averaged over a certain HR value and plotted as the graph shown on Figure 5 for absolute data and on Figure 7 for the normalized data. The Pd values were normalized using the Pupil Dilation Min/Max values ($BL_{min} = 1.34 \text{ mm}; BL_{max} = 6.15 \text{ mm}$) obtained from the baseline prestudy, using the normalizing formula:

$$ PD_{norm} = \left( \frac{DP_t - DP_{min}}{DP_{max} - DP_{min}} \right) \times (BL_{max} - BL_{min}) + BL_{min} $$

Considering the frequency distribution of the HR values ($N = 518077; \mu = 84.00, \sigma = 12.064, \sigma^2 = 145.532$), we can say that 95% of the responses fall between 61 and 110 of HR values. This Pd in relation to HR can be seen in the plot on Figure 6. This gives us a statistical significant range of values which exclude extremes and on which we can observe the pattern of behavior between pupil dilation and heart rate. Moreover, we can observe peak of dilation and the start of constriction of the pupil at the moment of 100 bpm.

9.5 Discussion

Our results showed a sinusoidal relationship between Pd and HR for the statistically significant range of values (extremes taken out) between 61 bpm and 110 bpm. This is reflected on both absolute and normalized Pd values (Figure 5 and Figure 7). This suggests that the Pd could be useable as a measure of analyzing the optimal arousal point, however only coupled with the HR because of the non-monotonous behavior of such a Pd function model. This has been pointed out in previously published work where it was stated that the pupil response will show an arousal effect only when the cognitive demands of the situation are minimal (Stanners et al., 1979; Steinhauser et al., 2004).

Previous published work found that the increasing task difficulty increases the HR (Boutcher and Boutcher, 2006; Sosnowski et al., 2004) and the Pd (Moresi et al., 2008; Recarte and Nunes, 2003). The increase in HR and the Pd can only be explained as reflecting physiological arousal (Stanners et al., 1979; Steinhauser et al., 2004; Bradley et al., 2008; Lang et al., 1993) This would support the finding in our study, at least up to (and not inclusive) 100 bpm point, where the pupil constriction starts. Since previous studies have reported that emotional arousal in Pd is associated with increased sympathetic activity (Bradley et al., 2008),
CHAPTER 9. IDENTIFYING THE MOMENT WHEN DEMANDING COGNITIVE LOAD OVERSHADOWS THE AROUSAL EFFECT ON PUPIL DILATION DURING A STRESSFUL TASK

Figure 9.5: The absolute pupil dilation responses averaged for specific Heart Rate values
Figure 9.6: The sample frequencies of the Heart Rate values
CHAPTER 9. IDENTIFYING THE MOMENT WHEN DEMANDING COGNITIVE LOAD OVERSHADOWS THE AROUSAL EFFECT ON PUPIL DILATION DURING A STRESSFUL TASK

Figure 9.7: The normalized pupil dilation responses averaged for the certain Heart Rate values
our results show that at the 100 bpm point the parasympathetic pathway activity dominates over sympathetic one and the constriction occurs. In this moment we could state that the cognitive load becomes too demanding overshadowing the arousal effect on the Pd. Just looking at the Hr, observing the activation of both parasympathetic (PNS) and sympathetic (SNS) pathways might prove a difficult task as the heart rate monotonously increases with task difficulty and arousal; thus, we cannot be sure in which moment one pathway activation prevails over the other. However, we show that by coupling Hr and Pd data such a moment emerges from the data and is identifiable as a peak of Pd data at the Hr of around 100 bpm.

We were able to observe (Figure 8) the point of optimal arousal according to the hypothesis stated in regards to the Yerkes-Dodson Law (Yerkes and Dodson, 1908). Interesting observation is that that point occurred on lower heart rate values for the normalized pupil dilation in this particular task.

9.6 Conclusions

This data coupling gives rise to another side of the analysis of psychophysiological arousal and the cognitive demands/difficulty of the task where demanding cognitive load overshadows the arousal effect; moreover, one can observe sympathetic and parasympathetic pathway activation and dominance over each other in those conditions. This can raise us another step towards pinpointing the activation underlying central nervous system mechanisms that contribute to cognitive activities.

9.7 Future Work

Wu et al. (2010) state that in an affective computing system which aims to improve the user’s performance in mental tasks such as learning it is very important to be able to identify the user’s optimal arousal level and to recognize whether or not the user’s actual arousal level is close to that optimal level. This experiment could be tailored to identify individual subject’s optimal arousal level in an attempt to personally customize this serious game task.
CHAPTER 9. IDENTIFYING THE MOMENT WHEN DEMANDING COGNITIVE LOAD OVERSHADOWS THE AROUSAL EFFECT ON PUPIL DILATION DURING A STRESSFUL TASK

Figure 9.8: The absolute (black) and normalized (red) pupil dilation responses averaged for the certain Heart Rate values
9.8 Acknowledgements

The research was carried out and funded as part of the xDelia research project (www.xdelia.org). We gratefully acknowledge funding from the European Commission under the 7th Framework Programme, Grant No. 231830.
CHAPTER 9. IDENTIFYING THE MOMENT WHEN DEMANDING COGNITIVE LOAD OVERSHADOWS THE AROUSAL EFFECT ON PUPIL DILATION DURING A STRESSFUL TASK
References


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


112


decision making using psychophysiological sensor data. In 51st Annual Meeting of Society for Psychophysiological Res.


REFERENCES


ABSTRACT

Emotions are thought to be a key factor that critically influences human decision-making. Emotion regulation can help to mitigate emotion related decision biases and eventually lead to a better decision performance. Serious games emerged as a new angle introducing technological methods to learning emotion regulation, where meaningful biofeedback information displays player’s emotional state.

This thesis investigates emotions and the effect of emotion regulation on decision performance. Furthermore, it explores design and evaluation methods for creating serious games where emotion regulation can be learned and practiced.

The scope of this thesis was limited to serious games for emotion regulation training using psychophysiological methods to communicate user’s affective information. Using the psychophysiological methods, emotions and their underlying neural mechanism have been explored. Through design and evaluation of serious games using those methods, effects of emotion regulation have been investigated where decision performance has been measured and analyzed. The proposed metrics for designing and evaluating such affective serious games have been exhaustively evaluated. The research methods used in this thesis were based on both quantitative and qualitative aspects, with true experiment and evaluation research, respectively.

Serious games approach to emotion regulation was investigated. The results suggested that two different emotion regulation strategies, suppression and cognitive reappraisal, are optimal for different decision tasks contexts. With careful design methods, valid serious games for training those different strategies could be produced. Moreover, using psychophysiological methods, underlying emotion neural mechanism could be mapped to provide optimal level of arousal for a certain task.

The results suggest that it is possible to design and develop serious game applications that provide helpful learning environment where decision makers could practice emotion regulation and subsequently improve their decision making.