A SERIOUS GAME FOR TRAINING IN EMOTION REGULATION
FROM DESIGN TO EVALUATION

Olle Hilborn
A Serious Game for Training in Emotion Regulation
From Design to Evaluation

Olle Hilborn
A Serious Game for Training in Emotion Regulation
From Design to Evaluation

Olle Hilborn
Licentiate Dissertation in Game Development

Department of Creative Technologies
Blekinge Institute of Technology
SWEDEN
A Serious Game for Training in Emotion Regulation: From Design to Evaluation

Abstract
Games are often used as training devices in various tasks, but proper biofeedback is more seldom used. Within an EU project it was explored how biofeedback games can target emotion regulation and be evaluated meaningfully. While many use games and biofeedback separately, here the focus was to combine them. This was explored through how the games were perceived and played while players were punished in-game, based on their physiological activity. By implementing games and study the interaction patterns in experimental settings, primarily correlational data was acquired. The results suggest that targeting cognitive constructs has to be validated for each specific game, since game strategies can influence the activation of the cognitive constructs.
Preface

Papers included in the thesis:


Related scientific work not included in the thesis:


A Serious Game for Training in Emotion Regulation: From Design to Evaluation ......1

Abstract ............................................................................................................................1
Preface ..............................................................................................................................2
Introduction ......................................................................................................................4
Games ................................................................................................................................6
Main concepts ................................................................................................................10
Research approach ........................................................................................................14
Contributions ................................................................................................................18
Conclusions ....................................................................................................................20
Future work .....................................................................................................................20
References .....................................................................................................................21

Paper 1..............................................................................................................................27
Impulsivity test implemented with game elements – can validity be maintained? ..............28

Paper 2..............................................................................................................................35
The Aiming Game: Using a Game with Biofeedback for Training in Emotion Regulation ..........36

Paper 3..............................................................................................................................58
A Serious Game Using Physiological Interfaces for Emotion Regulation Training in the Context of Financial Decision-making ..............................................................59

Paper 4..............................................................................................................................78
A study of the effect of emotion regulation in a shooting game using biofeedback from electroencephalography for training in emotion regulation........79

Paper 5..............................................................................................................................97
A biofeedback game for training arousal regulation during a stressful task: the Space Investor .................................................................98
Introduction
Games that are used for other purposes than entertainment are often referred to as serious games. Serious games have wide implication areas, such as education (Annetta, 2010), intercultural communication (Wiggins, 2011), and conflict resolution (Powers & Kirkpatrick, 2012). Biofeedback is one approach within the scope of serious games to teach people various aspects, such as better physical exercise (Goldstein, Ross, & Brady, 1977), lowering high blood pressure (Glasgow, Gaarder, & Engel, 1982), and pain perception (deCharms et al., 2005). Games have been developed with the use of biofeedback (e.g. Dekker & Champion, 2007; Kalyn, Mandryk, & Nacke, 2011; Kuikkaniemi et al., 2010; Rani, Sarkar, & Liu, 2005), serious games that make use of biofeedback are more rare, although they exists (Heiden & Lajoie, 2009). Problems that arise when biofeedback is introduced within game design are - among other - how to evaluate them; if they train what they are supposed to; and regarding how fun they are. One of the main issues in this work is the evaluation of the games. The other main issue is to implement a physiological concept that has a solid theoretical basis, but in a game setting with the specified purpose of offering teaching. It is important to validate that the setting does not affect our understanding of the phenomena.

Purpose
The purpose is to explore the relationships between games and cognitive constructs, as well as between games and emotion regulation.

xDELIA
The work reported in this thesis has been conducted within the EU-project Xcellence in Decision-making through Enhanced Learning in Immersive Applications (xDELIA). xDELIA was concerned with financial decisions made by private investors. Investors were defined by i) they trade their portfolio sufficiently often that systematic patterns are detectable, and ii) they trade on a regular basis through a trading platform (i.e. an on-line software application). Earlier training approaches for investors have used the tactic of increasing their understanding of the financial domain, but this has had little to no impact on trader performance. The approach used in xDELIA was to understand what comprises the investors’ expert performance and try to disassemble it into relevant components that can be trained.

Analyzing investors lead to the conclusion that investors may have adequate knowledge, but act on habits, emotional states, and attitudes that manifest as consistent decision biases. There is evidence that emotions may provide important
decision cues based on prior experience, helping direct attention in fast-paced decision environments (Pham, 2007). Thus the development of effective emotion regulation strategies which do not depend on simple suppression of emotions is likely to be an important element of expertise development within the financial domain (Fenton-O’Creevy, Nicholson, Soane, & William, 2005; Seo & Barrett, 2007). Many decision biases exhibited in laboratory studies turn out to disappear or have beneficial effects in naturalistic settings (Todd & Gigerenzer, 2007). However, some biases turn out to be remarkably robust in multiple contexts and have demonstrable negative impact on financial decision-making. Loss aversion (Kahneman & Tversky, 1984) is the tendency for people to be more aversive towards a loss of $X than they are attracted by a gain of $X. In other words, losses are more painful than the joy of an equal win. Loss aversion is one of the factors that gives rise to the disposition effect, the effect that investors (for instance) are more willing to recognize wins but prefer not to recognize losses, meaning that the net result from trading becomes less than zero under these circumstances (Camerer, 2000). Cognitive biases such as loss aversion and the disposition effect can be demonstrated not just in experimental settings but also in market data and can be shown to adversely affect financial outcomes (Camerer, 2000).

Recent research has started to address individuals’ susceptibility to biases and the role that emotion regulation processes play in this. For example, in a large scale qualitative field study it was shown that there were important differences in emotion regulation strategies between beginners and expert bank traders (Fenton-O’Creevy, Soane, Nicholson, & Willman, 2011). A recent laboratory study (Sokol-Hessner et al., 2009) showed loss aversion to be reduced by adopting an intentional cognitive approach to emotion regulation. Directing subjects to adopt such a strategy resulted in lower skin conductance, interpreted as less arousal when losses were suffered and in lower behavioral loss aversion compared to a control group. The Iowa Gambling Task, a task designed to show how risky behavior can be dependent upon the functioning of emotions (Bechara, Damasio, Damasio, & Anderson, 1994), and consist of different stages at which performance can be measured. A laboratory study (Heilman, Crișan, Houser, Miclea, & Miu, 2010) found that risk aversion (another important component of the disposition effect) and performance in the prehunch/hunch stage in the Iowa Gambling task was reduced. Performance in the gambling task was greater in a group that received instructions to take an intentional cognitive approach to emotion regulation compared to a control group.
Based upon this research in neuropsychology, and using wearable sensors with games, it was hypothesized within the xDELIA project that new game applications relying on biofeedback, rather than propositional knowledge, could be developed to support training in emotion regulation. Several games (Aiming game, Auction game, and Space Investor, described in detail later) were therefore designed with the aim of improving decision performance by

i) Making players more aware of their emotional states during relevant decision making tasks.

ii) Supporting learning and improving emotion regulation strategies to minimize the negative effects of emotions on decisions.

In other words, players received cognitive feedback, i.e. the return of a measure of the player’s cognitive process (on-line feedback), during gameplay, which has been shown to be superior to outcome feedback, i.e. once a task is completed the measurement is returned (off-line feedback), with regards to the content learned (Balzer, Doherty, & O’Connor, 1989).

**Games**

There are three games that have been developed with relation to emotion regulation and another with respect to impulsiveness. They are shortly described here.

**Punishment as learning**

Giving people a target value for their biofeedback data provides them with cognitive feedback on their performance. This has been shown to be able to enhance performance in a wide variety of fields, such as pain relief (deCharms et al., 2005), hypertension (Achmon, Granek, Golomb, & Hart, 1989), and high blood pressure (Glasgow et al., 1982). Not only feedback of the target value is given by the games described here, the actual difficulty of the games will change depending on how far away from the target value the player is. This results in that the players will receive feedback as well as a harder game experience when not performing regulation of the physiological signals. To achieve a good result in the game, the player has to establish a strategy combining both the regulation of physiological signals and specific game strategies, thus promoting regulation of physiology during gameplay tasks.
Aiming Game
The Aiming Game is a two-dimensional shooting game where the task is to shoot down as many black airplanes while avoiding shooting down red ones (see Figure 1). Each shot has a point cost in order to create a balance between shooting and aiming. The difficulty is raised by making the airplanes blurry and the aiming crosshair to move around. The Aiming Game design is described in Paper 2 and the experiments conducted are described in Paper 4.

Figure 1. Aiming Game in action. The black airplanes are targets while the red ones are distractors. Arousal information is given to the lower left.
**Auction Game**

The Auction Game is a financial decision scenario, where the player sees three price estimations and then decides if the right decision is to sell or buy (see Figure 2). The difficulty is raised by making price estimations harder to calculate by increasing the range of the estimations. It is described and used in the study in Paper 3.

![Auction Game](image_url)

*Figure 2. The Auction Game with the three price estimates on the screen. The arousal meter can be seen in the top right corner.*
Space Investor
Space Investor is a further development of the Aiming Game, but set in a 3d environment in space. The name comes from the intended audience (investors) and the setting (space), with the added pun from the classical game “Space Invaders” by Tomohiro Nishikado. The task is still to aim and shoot, but this time the targets are asteroids (see Figure 3). The game difficulty is raised by moving the crosshair, blurring the vision, increasing the speed of the spaceship, and weakening the shields (more damage when hit by asteroids). The spaceship is constantly moving forward, thus inducing a stressful element by having the asteroids approach the player constantly. In Paper 5 the Space Investor is described and tested.

Figure 3. The Space Investor. Arousal is represented by the 1/5 full green bar in the middle.
Main concepts

Impulsiveness

Before xDELIA was narrowed down to cover only investors, it was hypothesized that aspects of young adults’ financial behavior are based on impulsiveness. Paper 1 concerns itself with this, using the widely used definition of impulsiveness (Daruna & Barnes, 1993): “The behavioural universe thought to reflect impulsivity encompasses actions that appear poorly conceived, prematurely expressed, unduly risky, or inappropriate to the situation and that often result in undesirable consequences.”

Since impulsiveness is so multifaceted it is hard, and gives a faulty picture, to measure it as a whole construct. There are several models of how to divide impulsiveness into sub factors. In this work the division will be made according to a well-known scale, the Barrat Impulsiveness scale (BIS) (Patton, Stanford, & Barratt, 1995). BIS consists of six first-order factors:

- **Attention**, focusing at a task at hand (e.g. I concentrate easily)
- **Motor impulsiveness**, acting on the spur of the moment (e.g. I do things without thinking)
- **Self-control**, planning and thinking carefully (e.g. I plan trips well ahead of time)
- **Cognitive complexity**, enjoying challenging mental tasks (e.g. I like to think about complex problems)
- **Perseverance**, the ability to stay focused on a task (e.g. I can only think about one thing at a time)
- **Cognitive instability**, thought insertions and racing thoughts (e.g. I often have extraneous thoughts when thinking)

These factors can also be combined into second-order factors, where attention and cognitive instability are combined into attentional impulsiveness; motor impulsiveness and perseverance are combined into motor impulsiveness; self-control and cognitive complexity are combined into non-planning impulsiveness. These second-order factors were not used in the work presented here, due to the information they contain being available in the first-order factors.

Another measure of impulsiveness is the go/nogo task, in which a person is receiving continuous stimuli and must make binary choices (press or no-press). Reaction time and accuracy are collected for each response. Typically there are more go-trials than nogo-trials. There are several variations of this task including faces and circles (Schulz et al., 2007), numbers (Gomez, Cooper, & Gomez, 2000), and colored arrows (Vocat, Pourtois, & Vuilleumier, 2008), where the participant is
required to respond to one form of stimulus rather than the other. An impulsive person would have more false alarms (or incorrect responses) than a non-impulsive person, but no absolute levels exist, since individuals are compared to themselves and performance also depends on how the task is tuned.

**Emotions**
Emotions can be interpreted as manifestations of independent components of arousal and valence (Russel, 1980), where arousal represents excitement level and valence refers to pleasurable or unpleasurable feelings. Russell proposes that these independent dimensions represent two different underlying neurophysiological systems. The subjective experiences of specific emotions, such as fear or happiness, may be understood as consequences of cognitive interpretations of these patterns of physiological activity that occur in the context of eliciting stimuli (Posner, Russell, & Peterson, 2005). Hence, by this interpretation, emotions can be visualized in a diagram where arousal and valence define each axis, respectively, as seen in Figure 4. What are interpreted as discrete emotions can be measured by measuring a combination of valence and arousal.

![Figure 4. Emotions in the valence-arousal space.](image)
Since arousal and valence are two distinct systems, there should be different physiological correlates for them. Arousal has been reported to correlate with skin conductance and heart rate (Lang, Greenwald, Bradley, & Hamm, 1993), functional magnetic resonance imaging (fMRI) activation (Bradley et al., 2003), and electroencephalography (EEG) activity (Keil et al., 2001). Valence is correlated primarily with the activation of corrugator musculature (negative emotions) and zygomatic musculature (positive emotions), measured by measurement of skin surface voltage levels using electromyography (EMG, Cacioppo, Petty, Losch, & Kim, 1986; Lang et al., 1993). The work described in this thesis has used the EEG sensor EPOC developed by Emotiv\(^1\) in the Aiming Game (described further down), which is a commercially developed system that is meant to be used with games. Heart rate has been measured using a Movisens EKG, and arousal calculated with the software xAffect (Schaaff, Müller, Kirst, & Heuer, 2012).

**Sensor technology**

The sensors used are the Emotiv Epoc and the Movisens EKG. The EPOC was chosen while the development of Movisens EKG was taking place and the choice criteria were to have a system that could be used not only in a lab setting but in the investors’ real-life setting. Since the EPOC is a commercial EEG system built for games, it was considered a good choice, especially since an application programming interface (API) was available.

The EPOC is a 14 electrode EEG headset that connects wirelessly to the operating computer. It comes with a number of detection suites (Expressiv, Affectiv, and Cognitiv) that handle different interpretations of the data. What has been used for the work reported here is the “instantaneous excitement” variable from the Expressiv suite, which tracks arousal of a person in real-time. Arousal is represented as a value between 0.0 and 1.0, the value being passed to whatever game is being played. The EPOC also has other features such as a gyroscope and a crude EMG, but since these were not used, they will not be further described here.

The Movisens EKG is an electrocardiograph (ECG) device that measures heart rate. It is relatively non-intrusive since it is only 62x38x11mm, with a mass of 23g and can measure heart beats with a sample rate of 1024 Hz\(^2\). The Movisens device is connect using Bluetooth to the game computers and the data was preprocessed with xAffect (Schaaff et al., 2012) to give an arousal value between 0.0 and 0.99.

---


There was a requirement of a five minute baseline data collection period where the participants had to sit still and relax in order to zero the arousal level accordingly. The Movisens device also has acceleration and barometric sensors, but these have not been used in this work and will not be further expanded on.

**Emotion regulation**
Since emotions can affect the cognitive resources available to perform various tasks, it is important for task performance how emotions are handled. Antecedent-focused and response-focused emotion regulation are two broad categories suggested in the process model of emotion regulation (see Figure 5), developed by Gross (2002). Antecedent-focused strategies focus on regulating the emotion before the emotion eliciting event has happened, while the response-focused change the reaction to the event. These two categories are further broken down into:

- **Situation selection**, putting oneself in situations that are comfortable and non-upsetting, e.g. a person who is suffering from arachnophobia (fear of spiders) might not go into gardens for the specific reason of avoiding spiders.

- **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 garden she might run across it in order to minimize the time spent in it.

- **Attentional deployment** is to direct where the situation attention is placed. E.g. when the arachnophobiac is in the garden she might try to remember species of trees in order to turn attention away from potential spiders.

- **Cognitive change** is about selecting the meaning of the attended aspect of the situation, such as instead of thinking “I’m going to die by a spider attacking me”, thinking instead “no spiders in Sweden are actually venomous and they cannot hurt me”.

- **Response modulation** is the effort to affect emotions after they have already been elicited (response-focused), e.g. trying to hide the stress emotions that occur when seeing a spider.

Since in all experiments the participants are in the same situation and cannot change very much, the two categories relevant in this work are cognitive reappraisal (henceforth referred to just as **reappraisal**, mentioned as cognitive
change in the list above), 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 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 (Gross, 1998; Richards & Gross, 2000).

Wallace, Edwards, Shull, & Finch (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, 1998).

**Research approach**

There are various relationships that can be constructed by collecting data from the games. The most commonly sought for in science is causation, which establishes a cause-and-effect relationship. Not surprisingly, this can be very hard to find and give evidence for. Another common relationship is correlation, which means that two variables co-change. Correlation relationships are often described by how
strongly they correlate; even if a correlation is 100 %, this does not establish causation! The research here have mostly had a correlational approach, since varying many factors (such as in a biofeedback loop) makes it difficult to infer causal relationships.

**Research Questions**

The research has been directed by the research questions below. Under the questions are descriptions of how they have been approached in order to be answered in this work.

**RQ 1. Can games that utilize biofeedback be evaluated with regards to intended use/affect, design and level of difficulty?**

Several different analyses have been done with regards to the design of the games in this work, both qualitative and quantitative. These evaluations are of course not exhaustive (that would be theoretically impossible), but several variants were tried and are reported in each paper. The first papers are primarily about design while the latter ones also consider measures of difficulty in the games.

**RQ 2. Can the cognitive constructs impulsiveness and emotion regulation intentionally be engaged with game design?**

Cognitive constructs are hard to target with games, since even a slight change can drastically alter the results, as was shown in Paper 1 that discussed impulsiveness. However, emotion regulation, which was studied in Papers 2-5, is a more stable construct and much research has confirmed the validity of it. Also, there is indirect evidence for emotion regulation provided by psychophysiology sensors that can detect physiological correlates of emotional states. By tying the difficulty in the game to the player’s emotions, emotion regulation is coupled with performance in the game. Since all these games also require a degree of skill to play, one may argue that the skill in playing the game and the skill in regulating emotion cannot be separated. This is true, but only to an extent. The emotional values are intermixed with the performance, but the raw psychophysiological data was also analyzed. The question whether the players became better at the games because their emotion regulation capabilities were enhanced or if they showed reduced arousal due to their better abilities in the games is still an open question, but transfer studies are planned at partner universities that may provide evidence for answering this.
Methods

Experiment
The experimental method is a way to investigate causal relationships. By using the method of difference, i.e. to keep all things constant between two (or more) groups but the one(s) of interest, it can be possible to find evidence for causal relationships (Christensen, 1997). These relationships are determined by using statistical methods, which in this work consists of t-tests and ANOVAs. Following the standards in the field, a p-value of 0.05 was chosen, which means that there is a 5% risk that a difference identified between the groups, and thus conclusions based on this, is a consequence of random error.

How far the conclusions can be translated from a lab setting to the real world, also known as ecological validity, is a common issue for the experimental method. The work here has this question as well, but the aim has not been to evaluate how these games work in real life, but to outline different constructs that can later be tested in more naturalistic settings by xDELIA partners. The experimental approach was used in Paper 1, as well as in Paper 4, although that time unsuccessfully.

Cross-correlation
Correlation analyses provide tools for searching for relationships between test instruments without manipulating groups. They are also good when it is not clear what a control group would look like, such as the case with biofeedback games. Would a control group be playing the same game without biofeedback? Then the task would be completely different since emotion regulation would not be a part of it. If two groups for instance would have the same instructions, but one uses biofeedback while the other is not, the two versions of the game would differ a lot in their interactivity, which in turn would affect how the participants regulate their emotions. Would it be to lie to the participants and use a prerecorded difficulty setting, while the subjects think they play with biofeedback plugged in? Then the factor of trying to regulate the emotions, but not succeeding, is introduced in the picture.

The two versions of correlation that were used in the experiments reported here are Pearson’s and Spearman’s, since it was not possible to know if the relationships found would be linear or non-linear. Pearson’s product-moment correlation coefficient describes the strength of a linear relationship between two variables within the sample. A Spearman rank correlation coefficient only controls for if while one variable is growing, the other variable is either constantly shrinking or
growing (Walpole, Myers, Myers, & Ye, 2012). E.g. if the function \( f(x) = x^2 \) were to be tested, Spearman would successfully detect the correlation, while Pearson’s would have underestimated it. Another example is if \( f(x) = kx + m^2 \), where \( k = 0.5 \) and \( m = -1 \). In this case Pearson’s would give a very good correlation, but Spearman would not detect it.

**Data Collection**

Questionnaires

Much data about the participants has been obtained using questionnaires, which are widely used within various fields that deal with humans (e.g. psychology, informatics, human-computer interaction, cognitive science, etc.). The questionnaires used in this work have overwhelmingly used closed-ended questions when trying to pinpoint the various underlying constructs, since this is a good approach when answers are to be easy to answer, reduce variability and are easily coded (Bryman, 2012). However, in order to improve the design, some open-ended questions were also included in order to get creative input and catch things not covered by the close-ended questions.

The reasons for choosing questionnaires in these studies include minimization of interviewer effects, quicker and cheaper to administer, as well as keeping the experiments as standardized as possible (Bryman, 2012). All questions were designed to be at such a high level of measurement as possible, using only ordinal and nominal scales when necessary. For instance, age was always asked for in specific numbers, instead of an age range.

Some standardized questionnaires have been used during this research. The reason is that they have been validated previously and thus are more reliable than newly created ones. The specific questionnaires used were the emotion regulation questionnaire (ERQ) (Gross & John, 2003), the system usability scale (SUS) (Brooke, 1996), the Barratt impulsiveness scale (BIS) (Stanford et al., 2009), and the game experience questionnaire (GEQ) (IJsselsteijn, Poels, & De Kort, 2008).

Game data

Gameplay recording was event-based with time-stamps, and continuous recordings of the arousal data were made from the sensor connected to the game. All types of events had their own classification with a specific timestamp.
Contributions
The main contributions are various ways of making usability assessment; designs for biofeedback games; problematization of the go/nogo task and the connections to cognitive constructs; emotion regulation impact on game performance that goes against the hypothesis; and game assessment in relation to difficulty.

Each paper is listed below, together with the implications for the research. In addition to this, the contributions of Olle Hilborn (OH) are listed for each paper, since this has been recommended as good practice regarding authorship (Eriksson & Helgesson, 2013).

Paper 1
The topic of impulsiveness was explored and the go/nogo paradigm was used in order to assess impulsiveness. Four versions, the standard version and augmented versions with either music or animation, or both, of the go/nogo task were implemented. These versions were then cross-correlated with the BIS (described under “Impulsiveness”). The results implied that music did not impact the task, but animation and no-animation did. However, animation and no-animation correlated with different impulsiveness constructs, as measured with the BIS, suggesting that they both measure impulsiveness, but various kinds. There were also significant differences in the number of errors made by experiment participants.

As the main author, OH planned and revised the experiment, conducted the experiment, analyzed the results and wrote the manuscript. OH also revised it after comments from the others.

Paper 2
The idea to punish people when getting too excited was first introduced in the xDELIA project in the Aiming Game, where the EPOC sensor was used. The paper outlines the design, design decisions, software structure, and evaluation of the game. The evaluation consisted of continuous testing by the partner universities, a heuristic evaluation done by researchers at BTH, as well as a play testing with participants. The result suggested that the playability and usability of the game was good, although it is of course far from AAA game standard.

As the second author, OH was the primary discussion partner regarding design decisions during implementation, particularly regarding data acquisition. Moreover, early testing as well as the second largest writing task was conducted by OH.
Paper 3
Continuing with the work started with the Aiming Game, the context was set in a financial situation. A new sensor was also used, the Movisens EKG, an ECG device that measured arousal levels by heart rate. The evaluation results suggested adequate usability, although the arousal indicator was not attended to. Instead, other game elements gave players information about arousal. The small sample was difficult to interpret with statistical methods, and only one of the five arousal settings was distinct from the others with regards to game score.

As a contributing author, OH’s involvement was primarily proof-reading (although not agreeing to the statistics used) and suggesting important references that were missing, but also during the initial stage when the game was designed and theorized impact was settled.

Paper 4
In order to test for effects of emotion regulation tendencies, a full-scaled experiment was conducted with the Aiming Game. The results suggested that people with suppression tendencies actually performed better than reappraisers, something that went against the original hypotheses. In addition to this, it was found, to no one’s surprise, that experienced players performed better than novices. It was also found that novice players were not very affected by what emotion regulation tendency they had, while suppression correlated with good performance among the experienced players. It is hypothesized that once a minimum level of proficiency is achieved with the shooting aspect of the game, emotion regulation capabilities start to matter.

As the main author, OH designed of the experiment, ran pilot tests, analyzed the results, and wrote and revised the paper.

Paper 5
A new version of the Aiming Game was designed and implemented, and the game was renamed Space Investor (chosen because of the space setting in the game combined with the user domain: investors). The heuristic results from the Aiming Game study were considered during the design. Usability was maintained, if not improved, according to the play testing of Space Investor. Various levels were implemented and the results from testing suggested that the optimal strategy in the game was to regulate emotions, rather than learning to aim and shoot better. It was also apparent that lower arousal yielded an easier game and higher arousal
made the game harder, something that could not be conclusively shown in paper three.

As the main author, OH implemented large parts of the game, analyzed the results, and wrote and revised the paper.

**Conclusions**

The conclusions will be sorted under each research question.

*RQ 1. Can games that utilize biofeedback be evaluated with regards to intended use/affect, design and level of difficulty?*

Through questionnaires (GEQ and SUS) and play testing the design of the games has been evaluated, which resulted in much design input. In addition to this difficulty has been evaluated with performance measures from the game data. This data was compared to emotion regulation (in the Aiming Game) and with various arousal levels (in the Space Investor), something that yielded good data on difficulty levels.

*RQ 2. Can the cognitive constructs impulsiveness and emotion regulation intentionally be engaged with game design?*

Correlational data between game features and cognitive constructs have been shown. This is interpreted as a possible causal link between the game design and the constructs that can be further investigated. Moreover, the strategies to succeed in the Aiming Game and Space Investor games are primarily dependent on controlling emotions and not just to be good at aiming. These together suggest a good future training platform for emotion regulation training, although the pedagogical approach around such a platform is out of scope of the work here. The results regarding impulsiveness further strengthen the idea that games must be tested and validated when they are changed, since features that are not central to the core game still affected the validity.

**Future work**

Partner universities are trying to establish a causal link between playing the games and learning emotion regulation strategy by studying longitudinal effects of playing the various games. Tweaks in the games are yet to be explored in order to conclude an optimum learning level.
Game events presented here were evaluated with a success/fail distinction, but strategies within the games have not been addressed. This is natural because of the simplicity within the games, but could be an interesting direction in future research.

The paradigm used here is only concerned with immediate actions and input from the player. Although guided by psychological theories, no working model of the player is implemented. Future work would include creating a more dynamic view of the player that can help to change the game features not only momentarily, but during a longer period of time. One direction this could expand into is to start with player modeling from psychophysiological sensors. This could of course also be combined with behavior data.

References


Paper 1
Abstract. Four variants of the go/no task, which is an impulsivity test, have been implemented as computer games. A total of forty subjects tested the variants. The different implementations were compared against the Barrat Impulsiveness Scale (another impulsivity test) in order to see if all of the games target the same impulsivity constructs. Results showed that the different games actually measured different constructs, hence validity was not maintained. However, the games have more focused individual validity, since they correlate with other constructs, just not the same constructs as the original go/nogo task.

Introduction
Much research and discussion in the serious games community focuses on the development of learning and subsequent assessment of learning effectiveness (e.g. Anetta, 2010; de Freitas & Jarvis, 2007; Prensky, 2001). Here a different approach was taken that started with a cognitive task that was tested through several variants of game implementations in order to assess if the task’s diagnostic value can be maintained through these variations. The aim of this research is to be able to build small games, or game components, based on cognitive assessment tests. If successful, flash games could be created that players enjoy playing that at the same time provide validated information about various behavior tendencies for use in treatment and education programs. The benefits of using games for this purpose include facilitating cognitive tests in ways that can be distributed by the internet or packed with application software, and to provide built-in motivation and rewards for engaging with the tests when they are distributed electronically in this way. However, validation is critical, especially to ensure that gameplay does not interfere with the assessment functions of the cognitive tests.

The EU-project xDELIA (Xcellence in Decision-making through Enhanced Learning in Immersive Applications, Contract No. 231830) seeks to improve financial capabilities. One target group for this was young adults since this group prone to making financial mistakes (NSG-FC, 2009). In order to help this group, xDELIA has identified impulsiveness as a construct that is associated with negative financial behavior. Games that diagnose impulsiveness could thus increase awareness of how it bears upon life experiences, especially financial ones, and hopefully contribute to preventing unwanted financial behavior.

For impulsiveness, there is a commonly used self-report tool, the Barratt Impulsiveness Scale (BIS) (Patton, Stanford, & Barratt, 1995). BIS consists of six first-order factors: Attention, focusing on a task at hand; Motor
impulsiveness, acting on the spur of the moment; Self-control, planning and thinking carefully; Cognitive complexity, enjoying challenging mental tasks; Perseverance, the ability to stay focused on a task; Cognitive instability, thought insertions and racing thoughts.

The go/nogo task is a response-inhibition task, which means that the player is supposed to react as fast as possible in the response condition but must avoid reacting during the inhibition condition. The original task of the prototypes reported in this paper is modeled after the go/nogo test by Vocat, Pourtois, and Vuilleumier (2008). Measurements taken include reaction time (RT), omissions on go trials, and false alarms (FA) on nogo trials, where FA is seen as a measurement of impulsivity. The work presented here implemented go/nogo tasks with the added game elements of music and graphics (a moving robot character), based upon the hypothesis that: The assessment of impulsivity through the go/nogo task will remain the same even if specific game elements are added. This means that game elements, separately and in combination, had to be controlled and tested against the original impulsivity measures.

A recent model of work and performance is presented by Beal et al. (2005), which characterises peoples’ effectiveness on tasks in terms of how much cognitive resource they can allocate to them. A completely effective person engaged in a single task will allocate 100% of their cognitive resources to the task. This depends on the task attentional pull, off-task attentional demands (which can have an affective character, such as a feeling of sadness making thoughts wander) and the self-regulation and cognitive resources of the person (mediating the other two factors).

Posner et al. (1976) made four propositions about visual attention compared with attention to other sensory modalities: Proposition 1: Visual stimuli are not as automatically alerting as stimuli in other modalities. Proposition 2: In order for a visual event to serve as an effective alerting stimulus, the subject must first process the visual event by active attention. Proposition 3: The consequence of active attention towards any one modality is a reduction in the availability of the attentive mechanisms available for input to other modalities. Proposition 4: To compensate for the low alerting capability of visual signals, subjects exhibit a general attentional bias toward the visual modality whenever they are likely to receive reliable input from that modality.
In the study reported in this paper, Posner et al.’s propositions mean that in order for a player to react appropriately, the player needs to focus attention away from music (in the condition where it applies) towards visual input via a conscious action, since music should have a higher off-task attentional demand than vision, according to proposition 1. The player must keep a constant high attention level directed towards the game, focused on the visual modality, since the task attentional pull can be presumed to be relatively low in this case.

The music used in two of the prototypes was “The Murder” by Bernard Hermann, which has been shown to induce fear during guided visual imagery (Patton et al., 1995). Since emotions can affect performance, emotions were controlled for with the profile of mood states short form (SF-POMS; Shacham, 1983), which tests for depression, vigor, confusion, tension, anger, and fatigue.

**Design**

Four implementations of the go/nogo task were created: 1. as described in Vocat et al. (2008) called standard, 2. as the first implementation but with the music playing in the background (Psycho), 3. with an animated robot character running on a platform (LineRacer), and 4. with both the music and the animated character (LineRacer Psycho) (see table 1). In the LineRacer versions the players received feedback on their progress in the form of the speed of the robot character: a correct answer led to the character running (8 m/s), incorrect led to the character stopping (0 m/s), and no input led to the character walking at a medium pace (4 m/s).

<table>
<thead>
<tr>
<th>Character</th>
<th>No Music</th>
<th>Music</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Character</td>
<td>Standard</td>
<td>Psycho</td>
</tr>
<tr>
<td>Character</td>
<td>LineRacer</td>
<td>LineRacer Psycho</td>
</tr>
</tbody>
</table>

The players had to react to a certain stimulus; first, a black arrow was shown pointing either up or down, and then another arrow was shown. If the second arrow was green AND pointing in the same direction as the first arrow, the player was supposed to press the button "left ctrl" as fast as possible (go condition). If not, the player was supposed to not press (nogo condition). Gameplay consisted of two phases, the first one including 14 consecutive turns, and the second 33 consecutive turns. Both of these phases together completed one experiment cycle. The first
phase was a "calm" phase in which the player reacted as soon as possible, but without any other stressors (except for music in the relevant prototypes).

![Diagram of a turn and a cycle](image)

**Fig. 1. The parts of a turn (left), and a cycle (right)**

Each turn was divided into a relaxation period (500ms), a pre-response period (black arrow, random between 500-1000ms long), and a response period (green or blue arrow, 1500ms) (see figure 1, left). The animated character had a constant walking speed of 4.0 m/s. When a faulty answer was given, the character stopped for the rest of the response period. When a correct answer was given, the character started running (with a speed of 8.0 m/s) for the rest of the response period. The mean reaction time (on go-trials) was calculated for the 14 first turns in every other phase, then multiplied with 0.9, as in Vocat et al. (2008). This number, comparison number, was then compared with every new go-answer that was given. If the reaction time was higher than the comparison number, stress feedback was displayed telling the player that he or she can react faster. After 33 trials with stress feedback, a cycle was completed (see figure 1, right) and a new cycle began. Each cycle contained four nogo trials during phase one and eight during phase two. There were three cycles throughout the game, resulting in 30 go trials per participant without any stress and 75 potential go trials with stress feedback.

The score obtained by the player in the game was the distance the character travelled during the game. The score was calculated exactly the same way in the version where there is no animated character.

**Method**

Forty participants were recruited for the experiment, including students and people employed at the faculty. All participants were first asked to fill out a web questionnaire consisting of the BIS, the SF-POMS, and demographic data (age, gender, years of education, education program, and origin). Each participant was then randomly assigned to play only one of the four prototypes. After playing, all participants were again given the SF-POMS to detect if any mood changes had occurred during the experiment.
Results

Participants had a mean age of 28.7 (SD = 7.2), with 22 males and 18 females. Mean time spent in education was 15.75 years (SD = 2.3). Two people did not understand the task and have therefore been excluded from the results. Additionally, one person was feeling ill and sleepy and nodded off during the experiment and was thus also excluded from the results.

Analysis of variances (ANOVA) was performed, but no significant results between groups were found among the game factors of total score and reaction time. However, there was a difference in FA (false alarms, p = 0.028) between the groups. The Bonferroni analysis of FA did not reveal which of the groups had a big difference between them, so they were grouped based on character and no character, but also with music and non-music for further analysis (see table 1, grouping based on rows and columns). A t-test between the LineRacer (character) group (N = 17) and the non-LineRacer group (N = 20) showed a significantly larger amount of FA in the LineRacer group (p = 0.002), but no difference in reaction time or game score. Regarding the music group (N = 18) and non-music group (N = 19), no significant differences were found regarding reaction time, game score, and FA.

No differences were found between the groups (Standard, Psycho, LineRacer, and LineRacer psycho) regarding results from the BIS according to the ANOVA statistics. Another ANOVA showed no significant difference between the SF-POMS results before or after the game was played. T-tests within each group were performed in order to control for mood changes during the experiment; the only significant change was that LineRacer psycho felt less vigorous after the test (p = 0.017).

Table 2. Overview of Pearson’s correlations (Spearman’s rho marked where applicable) for total score, false alarms (FA), reaction time (RT), demographic data, BIS measures, and SF-POMS with regard to the groups LineRacer, non-LineRacer, Music, and Non-music. Significant correlations marked in bold.

<table>
<thead>
<tr>
<th></th>
<th>LineRacer</th>
<th>Non-LineRacer</th>
<th>Music</th>
<th>Non-music</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention - FA</td>
<td>-0.564</td>
<td>0.178</td>
<td>0.202</td>
<td>-0.368</td>
</tr>
<tr>
<td>Attention Total Score</td>
<td>-0.028</td>
<td>-0.499</td>
<td>-0.316</td>
<td>-0.211</td>
</tr>
<tr>
<td>Attention RT (Spearmans)</td>
<td>0.091</td>
<td>0.470</td>
<td>0.181</td>
<td>0.215</td>
</tr>
<tr>
<td>Motor - FA</td>
<td>-0.495</td>
<td>0.136</td>
<td>0.41</td>
<td>-0.332</td>
</tr>
<tr>
<td>Motor FA (Spearmans)</td>
<td>-0.606</td>
<td>-0.009</td>
<td>-0.156</td>
<td>-0.258</td>
</tr>
<tr>
<td>Cognitive Instability RT</td>
<td>-0.508</td>
<td>-0.033</td>
<td>-0.288</td>
<td>0.115</td>
</tr>
<tr>
<td>Cognitive Instability Total Score</td>
<td>-0.453</td>
<td>-0.22</td>
<td>0.016</td>
<td>-0.458</td>
</tr>
<tr>
<td>Age Total Score (Spearmans)</td>
<td>-0.190</td>
<td>-0.096</td>
<td>-0.041</td>
<td>-0.590</td>
</tr>
<tr>
<td>RT-FA</td>
<td>-0.181</td>
<td>-0.411</td>
<td>-0.616</td>
<td>-0.84</td>
</tr>
</tbody>
</table>
Correlation data for reaction time, total score, FA, demographic data, BIS, and SF-POMS are seen in table 2. No correlations were found for the regular groups (except interrelations within SF-POMS, which are already shown in Baker et al., 2002). Another finding (t-test) was that RT significantly differed \( (p = 0.006) \) between people who have an origin in western cultures (Sweden, Britain, and Australia) and in eastern cultures (Iran, Bangladesh, and Pakistan), where the westerners were faster.

5 Discussion and Conclusion

The hypothesis “The assessment of impulsivity through the go/nogo task will remain the same even if specific game elements are added.” is not supported by the results from the LineRacer and non-LineRacer comparison, since there was a very clear difference in the number of FA between these variants. Emotional reporting suggested no difference between these groups, suggesting that emotions did not increase off-task attention. Posner et al.’s (1976) second proposition that a visual event must be actively attended in order to be properly processed is a plausible explanation model; when a player has the running character, since it is moving and is the only way of getting feedback on how the player is doing in the game, it might be more interesting for the player to look at the character (constituting off-task attention pull), thus losing task focus.

The BIS constructs attention and motor impulsiveness were strongly negatively correlated with FA within the LineRacer group, so those games are a good candidate for further implementation as tests of impulsiveness. However, the task would no longer have the support of the current literature since the non-LineRacer group, which can be considered a more “pure” version of the go/nogo task, has opposite correlations on each impulsiveness factor in table 2 compared with the LineRacer group. As people scored higher in the attention construct, their RT increased (which explains the negative correlation between game score and attention, since a good score is dependent on low RT). This is a very peculiar relationship, since better attention is expected to speed up reaction, not slow it down.

There were no findings of a tradeoff between RT and FA, except for in the music group. This tradeoff is interesting since it was not found in the other group, nor in the literature on which these games were built (Vocat et al., 2008). It could be that the music induces this tradeoff.
The data presented above suggests that since variations of the task measured different constructs, the variations were no longer the same tasks. Even though most task characteristics were constant, the added animations and music recruit other cognitive systems or resources than the ones used in the standard task.

A concluding reflection regarding different go/nogo tasks is that it would be possible to try to implement a completely new task with much higher attentional pull right from the beginning, in order to prevent a standard example of the task from being polluted by distractors (game features). This way the task could be cross-validated with existing impulsiveness measures to accomplish a game assessment tool for impulsiveness. The reason for not taking this approach in this study was that a gradual increase of game elements was of interest in order to asses a derivation of the go/nogo task as an assessment tool for impulsivity in game form, and if the game elements did not have any influence upon cognitive processing, then it, in the game form, may serve no purpose.

References

Paper 2
The Aiming Game: Using a Game with Biofeedback for Training in Emotion Regulation
Henrik Cederholm, Olle Hilborn, Craig Lindley, Charlotte Sennersten, Jeanette Eriksson

ABSTRACT
This paper discusses the development of the Aiming Game, a serious game intended to be used as a tool for training emotion regulation. The game is part of an intervention package designed to support training of financial investors in becoming aware of their emotional states as well as providing them with a toolbox which can be used for training to counteract cognitive biases which may interfere with their trading activities. The paper discusses how such a game can be implemented as well as how it can be effectively evaluated. The evaluation is mostly focused on the effectiveness of the induction of emotional arousal by the game, which is supported by standardized game design methods and patterns.

INTRODUCTION
In the last few years, computer games have started to become valuable tools for different kinds of skill training Garris et al. (2002). These types of games, or serious games, can be designed very differently depending on the type of training they are intended to provide. Simulation games generally try to replicate a real life scenario, such as pilot training or stock trading, in order to give the player direct training and transferable skills. Some serious games, like the one described in this paper, aim at training a specific skill in a game setting, which in turn is hypothesized to be transferable into a real world setting. Using serious games in this manner has the obvious advantage of being both cheap and risk-free in comparison to practice in real life settings as well as specifically targeting specific cognitive skills and processes.

This paper describes a serious game called the Aiming Game (2D), which is used to train players in becoming aware of and controlling their own states of emotional and physiological arousal. The Aiming Game is a two-dimensional shooter game where the player tries to aim and shoot down airplanes, as shown in figure 1, while trying to regulate his or her emotional state in order to receive certain in-game advantages.

The game is part of an intervention package aiming at increasing performance in investment decision-making settings in the xDELIA project (Excellence in Decision-making through Enhanced Learning in Immersive Applications, Contract No. 231830). This paper will elaborate on the development process, design and evaluation of the game prototype. It described the process of the first development iteration cycle as well as the initiation of the second iteration based on the results from the first. It will not, however, provide solid evidence that transferable skill training is taking place since this study has yet to be performed by project partners.
Figure 6. The Aiming Game.

**Purpose**
The main concern of the xDELIA project and thus the Aiming Game is to develop learning interventions for financial investors, particularly those using the Saxo Bank trading platform. We focus mainly on investors who meet the following criteria:

1. They trade their portfolio sufficiently often that systemic patterns and biases in their trading are detectable.
2. They trade on a regular basis through a trading platform.

There is evidence to suggest that effective regulation of emotions can have positive effects on performance in investment and trading settings (Fenton-O’Creevy et al., 2010). The Aiming Game is specifically aimed at assisting investors in becoming aware of their own arousal state as well as training them in regulating their arousal. Successfully training investors in arousal regulation in a game environment is hypothesized to have a positive effect on their behavior, in a real trading environment.

**Emotions and Emotion Regulation**
Emotions can be interpreted as manifestations of the independent components of arousal and valence (Russell, 1980), where arousal represents excitement level and valence refers to pleasurable or unpleasurable feelings. It is hypothesized that these independent dimensions represent different underlying neurophysiological
systems. The subjective experiences of specific emotions, such as fear or happiness, may be understood as consequences of cognitive interpretations of these patterns of physiological activity that occur in the context of eliciting stimuli (Posner, Russell, and Peterson, 2005). Hence, by this interpretation, emotions can be visualized in a diagram where arousal and valence defines each axis, as seen in figure 2.

![Figure 7. Emotions in the valence-arousal space.](image)

By this interpretation, emotions can be measured by measuring a combination of valence and arousal. Valence has been inferred by facial electromyography (EMG) (Fridlund & Cacioppo, 1986), but since there are several technical difficulties such as accessibility and extensive setup procedures, the version of the Aiming Game described in this paper is not concerned with valence. Instead, the focus here is on arousal as the primary attribute of interest, with the game being developed as a tool for training in deliberate regulation of state of arousal.

When facing difficult and stressful tasks, people tend to use one of two main strategies to deal with corresponding negative emotions (Wallace et al., 2009). These strategies are:

- Suppression
- Reappraisal
Suppressers generally tend to push down emotions, but are continuously affected by them to a large extent. Reappraisers however tend to positively reevaluate situations, making the emotional response of the new perceived situation more bearable. Both emotion regulation strategies exhaust cognitive resources for the person affected by the emotion (Wallace et al., 2009). Wallace et al. point out that suppressing emotions generally takes up more cognitive resources than reappraisal. Generally it is therefore preferable to apply reappraisal strategies to deal with unwanted emotions.

In order to identify emotion regulation strategies used by individuals, Gross et al. (2003) developed the Emotion Regulation Questionnaire (ERQ). The ERQ makes specific statements in regards 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”. Results of the ERQ can be cross-correlated with results from a demanding task, such as the Aiming Game.

**DESIGN**

The Aiming Game is one piece in a learning intervention meant to facilitate people's, particularly investors who trade on a regular basis, learning to become aware of their emotional state. It is also meant to provide effective training in how to more efficiently regulate and control their emotions. While this training occurs in a game related setting, we hypothesize that the resulting skills that are learned can be transferred to the different settings, such as financial investment activities.

**Gameplay**

The Aiming Game is a two dimensional, first person shooter game, developed in Unity 3D™, where the main objective is to score as many points as possible by shooting down targets in the form of black airplanes. This is done by using a regular computer mouse as input device to aim at and shoot targets. The core game consists of three levels or phases, each lasting 180 seconds. The phases and their respective additional core game mechanics are explained in the table below.
Table 3. The three phases with objectives and features.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Objectives</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 1</td>
<td>Shoot down targets</td>
<td>Targets (Black airplanes)</td>
</tr>
<tr>
<td>Phase 2</td>
<td>Shoot down targets/Avoid distractors</td>
<td>Targets (Black airplanes)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Visual distractions (Red airplanes)</td>
</tr>
<tr>
<td>Phase 3</td>
<td>Shoot down targets/Avoid distractors</td>
<td>Targets (Black airplanes)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Visual distractions (Red airplanes)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Auditory distractions</td>
</tr>
</tbody>
</table>

The first phase is basically an introduction to the core game mechanics and a chance for players to learn how to play. The player attempts to shoot down airplanes as they appear from outside the screen and rapidly move across it. These targets are spawned once every 0.8 seconds.

In the second and third phases of play, visual distractions are added in the form of red airplanes. The goal is still to hit the black planes, but also to avoid shooting down the distractions. The purpose of the distractors is solely to disturb the player and pressure him or her into making errors, and thus becoming stressed. Distractors are spawned once every 0.4 seconds.

In the first version of the Aiming Game prototype the velocity and spawn frequency were exactly the same. According to an early heuristic evaluation of the prototype, related to game challenge as described in Isbister & Schaffer (2008), it was found that the element of distraction had too small an impact and was not challenging the players in a stimulating way (e.g. see Gee, 2005). Adjustments to the visual distractor were made accordingly which resulted in the red airplanes moving 30% faster than the black ones and also spawning 100% more often.
The third phase involves auditory distraction by adding stressful music. North & Hargreaves (2008) argue that music plays a role in task performance and showed that music and concurrent tasks competed for the same cognitive resources. In the Aiming Game the song Surfin’ Bird by the Thrashmen is used because of its stressful nature.

There is no limit to how many shots one can fire in a certain amount of time during the game. To ensure balance regardless of play styles when it comes to fire mechanics, a shot cost was implemented. This means that for every shot fired a score of two points are reduced from the player’s total score pool (compared to a ten point gain when hitting targets and a ten point loss when hitting distractors). Without this feature, it would become beneficial to shoot frantically without hesitation, consideration or strategy.

**Bio-Feedback and Game Play**

Biofeedback concerns the process of making users, or in this case players, aware of some physiological state in their own body. Today, biofeedback is widely used in, for instance, medical studies (Babu et al., 2007) to treat various kinds of pain and disorder. To provide biofeedback in the Aiming Game, the Emotiv EPOC™ (Figure 3) is used (http://www.emotiv.com). The EPOC provides Electroencephalography (EEG) and Electromyography (EMG) sensors that can detect electrical signals produced by the brain and facial muscles, respectively, that can be interpreted as measurements of the instantaneous excitement, or arousal, of a player.

![Figure 8. The Emotiv EPOC.](image-url)
This information is used as bio-feedback, meaning that it is fed back into the game and displayed in a bar on the screen, allowing the player to become aware of his or her current state of arousal. Arousal is divided into five segments, from one representing very low arousal (completely calm) to five representing very high arousal (highly excited/stressed).

Besides presenting the player with psychophysiological information, the data is also used to create distractions in the game depending on the player’s current arousal level, in two different ways:

- Distorted aiming
- Blurring of targets and distractors

Aiming is distorted by receiving an offset to the original aim position. This offset is constantly moving within the bounds of a predefined square (Figure 4) and the distance between the original position and the offset position is directly related to the amount of arousal one is experiencing at the time.

Also, when the player becomes aroused, the targets start to become blurred. The amount of blur affecting the airplanes is balanced so that, at minimum arousal, there no blur at all, while at maximum arousal it is hard to see the airplanes' exact position.

While constantly providing the player with real time bio-feedback and affecting game content depending on arousal, all game data is also logged to file to allow
retrospective analysis of both in-game actions and arousal. This allows the Aiming Game to be used as a tool for studies in addition to being a training environment for emotion regulation.

**Motivation**

An issue to consider when developing serious games is that there may not always exist intrinsic motivation for the player playing a game or to invest much time in it, since the motivation behind developing the game is not primarily entertainment. In the case of the Aiming Game, with the investor target group, it may be that the desire of players for learning and skill training will be motivation enough to get them to frequently use the platform or game.

Even though the Aiming Game is a relatively simple game with regard to game mechanics and features, it is vital that the elements that do exist both help and support the motivation to play the game. The main design goal of the game is for it to be challenging and thus also to allow players to practice mastery (Schell, 2008). Furthermore, Karat et al. (2000) claim that there can be great satisfaction in the ability to master one’s tools and produce a desired result, so users are willing to invest a great deal of time in doing so. Offering challenge and the opportunity to master a skill therefore seems to provide great, and perhaps even sufficient, motivation for people to engage in games.
Game Logic and Elements
The Aiming Game mainly consists of the software elements described in Figure 5.

The game elements mostly revolve around the Shooter object, where most actions and calculations are performed. This is where bullets are generated when the player shoots by clicking the mouse and also where the mouse object is updated.

With each frame of the simulation, the Shooter object calls the Arousal Controller to request updates of the psychophysiological data collected from the EPOC-wearing player. The response is a value between one and five which is then translated and sent to the mouse object where the aiming offset is applied accordingly. At maximum arousal level, the crosshair will receive an offset of approximately 15% of the screen width.

There are two object-generating entities called Target Spawner and Visual Distractor Spawner. These entities run on predefined timers and spawn (generate) their respective child objects (Targets and Distractors) according to their timer intervals. Targets are generated with a frequency of 1.25 (i.e. every 0.8 seconds), and Distractors with a frequency of 2.5 (i.e. every 0.4 seconds).
For each frame, the Collision detection object compares Bullet objects to both the Distractor objects and the Target objects in order to identify collisions, where a collision means that a bullet actually hit an airplane. The Collision detection object then calls the appropriate actions such as explosion animations, sounds and score adjustments.

Regardless of the rest of the scene, there is a timer object counting down from a predefined time, thus keeping track of when to change levels, as well as a Score Counter which, during all phases, collects all score data. Scores are calculated by the following criteria:

- Shot: -2 points
- Target hit: +10 points
- Distractor hit: -10 points

Connected to the Timer is also an Audio Player. This is a simple entity which becomes active in the third phase and controls the background music that is meant to distract the player.

**Data gathering and analysis**
The incentive for data collection in the Aiming Game prototype is twofold. The first and most important reason is to present relevant data to the player in real time. Malone (1982) stresses the importance of players always being able to identify their score or progress in the game. At the same time, the game interface should be as non-intrusive as possible so not to interfere with the player’s attention. Logged components that are also necessary to represent visually to the player during gameplay are:

- Real time arousal value
- Score

The second incentive to gather data is to be able to perform analysis of the participant performance regarding both score and success rate in emotion regulation. Data is therefore stored in two separate files, namely arousal statistics and shot statistics, for each phase and each participant. The data that needs to be collected for each phase is:

- Participant ID
- Phase
- Play time this phase (in seconds)
- Number of samples
- Sampling frequency
- Arousal value at incremental time points
- Shots and consequences and when they occur
- Total shots fired
- Total hits

The two data files are uniquely identified by Participant ID to ensure complete anonymity. The samples related to time are collected at each game frame to ensure that there is not a lack of data. Each sample contains a time stamp and an arousal value. The arousal value is stored in the text file with values that range from one to five, one being very low arousal and five being the maximum arousal that the EPOC device is able to register. In order to analyze changes in arousal over time with regard to in-game actions, we also store shot statistics for each action, including the time and outcome of that specific action. This can be used to investigate how temporary failures (misses and distractor hits) may be related to changes in arousal, by correlating information from arousal statistics and shot statistics.

**Evaluation**
In order to assess the potential of the Aiming Game, several evaluation methods were applied. The evaluation of the Aiming Game consists of studies related to game mechanics and usability testing using Heuristic evaluation as well as play testing by students, all parts of a generic Evaluation Toolkit developed in the xDELIA project. Usability refers to user interfaces and how helpful the game or system is in providing the player with necessary information and guidance, while gameplay heuristics analyze the actual gameplay and how well it is designed in terms of game content.

**Heuristic Evaluation**
The game development iteration was followed by a heuristic evaluation which aimed to qualitatively identify design errors and suggest improvements to correct them (Desurvire, 2004). The Heuristics are divided into a set of categories inspecting different aspects of the game prototype. There are several more Heuristics to the evaluator’s arsenal but here only the ones that actually had an impact on the evaluation of the Aiming Game are discussed. Heuristic evaluation
requires three evaluators and it is desirable that all three evaluators have both
game design and usability competence. The evaluators evaluate the game
separately and are not supposed to collaborate with each other. The list of the
heuristics is distributed to the evaluators after which the evaluators describe issues
violating each heuristic in the list. When the first step is completed, the evaluators
meet and put together their lists of issues. If two evaluators have the same issue on
their lists the problem stays on the final list. All the issues are discussed and if only
one has a particular issue and the others can agree that the issue is legitimate then
that problem stays on the list as well. A report is then prepared which describes the
issues in more detail than in the previous list. This report should contain
screenshots to clarify the issues more effectively. The evaluators together with the
developer discuss possible solutions and the suggestions are compiled into
recommendations and added to the documentation. After the heuristic evaluation
is conducted and documented, the result is presented to the product owner and in
collaboration decisions regarding what to do with the recommendations are made,
including whether a new iteration should start or whether play testing should be
conducted.

**Play Testing**

Play Testing is a valuable asset in the evaluation of games since it allows different
players to analyze the game from different subjective perspectives. This may reveal
novel aspects of experience which have not previously been discussed or evaluated
by the development team. Different components of game experience can be
measured using the Game Experience Questionnaire (GEQ) (IJsselsteijn et al.,
2008).

Additionally, while participants play the game, psycho-physiological data is
collected along with performance data. This data is retroactively compared with
the participants' subjective statements.

**Procedure**

For the Play Testing Evaluation, six arbitrarily chosen people played the game
(following the recommendation of Pernice & Nielsen, 2009) for approximately 10
minutes each. Before playing the game, demographic data considering gender, age
and experience with similar digital games was collected. In order to objectively
determine which game elements the players are paying most attention to, the
gameplay can be studied using an eye tracker device (as exemplified by Sennersten
2008, 2010). Eye Trackers measure the saccades (fast movements) and fixations
(dwell times) of human gaze (Duchowski, 2003) and have been used for several
decades for different assessments, e.g. Graf and Krueger (1989) and Lankford et al (1997). Graf and Krueger (1989) applied Eye Tracker technology to both measure usability and analyse Eye Tracking as an interaction device, while Lankford used the Eye Tracker to identify important graphical features in software design. Due to the relationship between eye fixations and attention focus, we are able to infer aspects of cognitive processes underlying virtual environment exploration (Sennersten, 2008). The purpose is to be able to tell how important different objects are, and in interviews afterwards, receive indications of how to improve the visual representations of the game.

Participants were seated in front of the Eye Tracker and were connected to the computer via the Emotiv EPOC. They were also briefed on the equipment and the purpose of the study. Instructions were given to the player from a static screen on the computer and were not communicated through the experimenter in order to provide each participant with the same preconditions. The participants were then asked to play the game and follow the instructions on the screen which led them seamlessly through the play testing session. After playing the game, participants were asked to complete the Game Experience Questionnaire (GEQ) (IJsselsteijn et al., 2008) as well as a modified version of the System Usability Scale (Brooke, 1996).

According to IJsselsteijn (2008) the GEQ measures experiential components of immersion, tension, competence, flow, negative and positive affect, and challenge. Each of these seven components is assessed by 5–6 question items (e.g., "I was deeply concentrated in the game" is a flow component item). Each question item consists of a statement on a five-point scale ranging from 0 (not agreeing with the statement) to 4 (completely agreeing with the statement).

Game Usability can be measured with a modified System Usability Scale (SUS) (Brooke, 1996). Brooke states that the SUS has proven to be a valuable evaluation tool, being robust and reliable. To evaluate results, Tullis & Albert (2008) argue that an average SUS score under 60% is relatively poor and one over 80% can be considered good.

In addition to the play testing questionnaires, semi-structured interviews were conducted in order to catch additional information about gameplay and the players’ experience. The players were questioned about specific game elements that caught their attention, if anything was lacking, motivational factors that were present (or missing), and game strategy. These were conducted after the
questionnaires, and after the player had enough time to reflect upon their gaming experience.

The interview questions usually became open discussions regarding the specific planned topics, which in turn generated additional useful data. The topics discussed with participants were:

- Arousal bar and Bio-feedback
- Aiming Mechanics
- Flow and Progression
- Difficulty
- Music
- Other (open discussion to let participant speak freely regarding the game)

Elements that the players were uncomfortable with and suggestions for gameplay improvements gathered in the play testing session were fed back into the design process as work items for the next development iteration.

**Play Testing Results and Discussion**

The Play Testing evaluation generated much valuable data, which was used in order to improve the Aiming Game both as an entertainment platform and as a learning tool.

When the results from the GEQ, SUS and interviews were collected and summarized, experimenters attempted to put the new-found information in the context of what it means in terms of game design.

An observation regarding the Arousal Bar was that five out of six participants claimed to have had awareness of their own arousal with the help of the arousal bar. Eye Tracking results however indicate that none of the participants paid any (or very little) direct attention at all to the arousal bar during gameplay. This finding might indicate that players are able to perceive the arousal bar in their peripheral vision while playing the game and having the centre of vision focused on the bar is not necessary for awareness of its value. Another possible explanation for the participants claiming that at most times they had full awareness of the level of the arousal bar might be that they received this information in other forms. This hypothesis is based on the fact that players receive several indications of their arousal in the form of airplane blurring and crosshair offsets as well their own gut feeling. This phenomenon will have to be investigated further.
In general participants answered uniformly with respect to the Aiming Mechanics. All participants experienced aiming as being somewhat rough and several participants drew the analogy to an old mouse with wheel-based mechanics.

Most participants (five out of six) also had a problem with the delay between when a shot is fired and when it actually hits the targets. Players had to learn how this worked before being able to hit targets correctly. Since the game does not explicitly explain the phenomenon, this caused confusion among participants.

Several participants described the development of tactics throughout the game and how these changed in accordance with the changes between game phases. The most common tactic seemed to be to focus attention towards the middle of the screen where evidently the most planes eventually appeared. This tactic was shattered however when the second phase was initiated and red distraction airplanes appeared. Since these come in greater number with the same generation procedure as the targets, the red airplanes swarm in the middle of the screen making it very hard for the player to separate the targets from the distractors in this region. This led to players completely switching tactics and focusing their attention to the borders of the screen instead, something that can be seen in the Eye Tracker data of some participants during the last phases.

When asked what their suggestion to make the game more interesting to play would be, all participants talked about the repetitiveness which eventually leads to boredom in the game and suggested different ways to vary the game content. Suggestions to decrease repetitiveness were:

- Additional types of airplanes with different features such as blue planes which are rewarded with more points.
- Reaching new levels by performance instead of time, to increase the incentive to achieve better.
- Several graphically different levels
- A variation of many games with the same mechanics but different game manifestations, e.g. click-and-drag items to different places.

From the Game Experience Questionnaire answers, it was possible to extract uniform trends regarding several components. In order to gain statistical significance from the GEQ one would preferably need approximately 20-30 participants. In the case of the Play Testing of the Aiming Game, which only used
six participants, this is not possible. In this section several components will be discussed which were nevertheless apparent and should be taken into consideration in further design iterations.

Participants generally answered that they felt tense during the game (Average: 4, Standard deviation: 0.89) as well as having to constantly focus on the game (Average: 4.17, Standard deviation: 0.41). They also congruently reported low scores on the questions of whether or not they were allowed to explore things (Average: 1.5, Standard deviation: 0.84) and if they felt imaginative (Average: 2, Standard deviation: 1.09).

Participants also reported low scores on whether they felt happy during the game (Average: 2, Standard deviation: 0.63) at the same time as stating that they were not particularly bored (Average: 1.83, Standard deviation: 1.17).

**Design Process**

The concept of the Aiming Game was developed in a collaborative manner similar to Participatory Design (http://cpsr.org/issues/pd/introInfo/) where partners met and sat down to discuss ways to induce player arousal in a relatively simple game. When the game product had been roughly defined, a product owner was appointed and developers began to specify their interpretation of the agreement on paper which was then reviewed by other project partners to find out if everyone's interpretations of the concept were congruent. When partners had agreed on the documented concept, the specification was broken down using a traditional Breakdown Structure (WBS) which resulted in a SCRUM (http://www.scrum.org/) backlog (requirement list). Since the backlog had to be approved as well, this was sent to the product owner as well as involved partners/stakeholders. Once approved, the development of the first iteration could begin. The development process proceeded for two weeks before developers and the product owner met again for a check-up meeting to make sure development of the game prototype was on track. Once confirmed, another two weeks of development followed.

When the "final version" of the first prototype iteration was completed, a Heuristic evaluation was applied followed by a Play Testing session. Heuristic evaluation can be very beneficial in such an early stage of development since it usually is able to detect design errors that developers might have missed. Play Testing, however, is more concerned with Game Experience which might be harder to assess so early in the iterative development process.
DISCUSSION
The Aiming Game was developed to assist investors in learning to identify and regulate their emotional state, more specifically their arousal level. In a first iteration of the development of such a learning platform, the game had to support designs which can be applied to a general audience, before specifically targeting the investor group. The blind implementation of the initial requirements was sure to lead to design faults. Being aware of this fact and having specific methods to tackle it was therefore crucial. Heuristic Evaluation, executed by colleagues, followed by a Play Testing session were applied in order to identify these design flaws and systematically structure and prioritize them. These methods generated a large quantity of faults, opinions and suggestions on how to improve the prototype. When using Heuristic Evaluation, we recommend analyzing the heuristics and eliminate irrelevant items. In the same way, one might realize that "standardized" heuristics do not cover all aspects of a game. In our case there were no heuristics for evaluating novel interaction devices, so we had to develop our own heuristics for these.

Interestingly there was not always a congruent philosophy between evaluators and play testers. Certain topics split the groups into two camps expressing completely different opinions and also suggesting very different solutions, while in other cases, issues expressed by one group was rejected as issues at all by the other. Both the Heuristic evaluation as well as Play Testing have been quite successful in their purpose of identifying design flaws and will therefore be applied in future to similar development scenarios.

During the Play Testing, and also during interviews with test participants, it became clear that it is important to bear in mind the potential difference between Play Testing subjects and the ultimate target group. In the case of the Aiming Game this difference was quite large since our Play Testers mostly consisted of students and colleagues while the target group is financial investors having very different backgrounds and experience. However, this difference does not exclude the possibility of receiving much helpful data from such a study. Regardless of gaming experience, all players should immediately understand the purpose, goal and mechanics of the Aiming Game. When game rules or other elements were unclear to subjects, this was immediately registered as a fault in the game design. In this way, the Aiming Game strives to become an intuitive tool suitable for anyone, regardless of experience.
One of the greater challenges with developing games that should suit players with very different levels of experience is the balance of difficulty. The Aiming Game should be tuned in such a way that it can be stimulating and meaningful to experienced players as well as people with limited knowledge and experience with computer games. The current design, having elapsed time as the only criterion for switching between phases, does not fully support this requirement. Instead, in ongoing development, the game should progress by other criteria, such as player performance.

From the Play Testing and interviews, it was obvious that for the Aiming Game to work as a long-term learning platform, it has to become more interesting for players. According to the results of Play Testing, players found the game to become quite boring even before having played an entire session (4x 3min). Since the idea is that the game should, at least to some extent, entertain players while they are also learning/training, the complexity of the game and the variety of game elements must be increased in order to keep players motivated.

In the Aiming Game the song Surfin’ Bird by the Thrashmen is used because of its stressful nature. The subjectively chosen music will be compared in a future study with music generally accepted to induce stress, first listed by (Mayer et al., 1995), to analyze if there is a difference in average stress levels between groups. Hints at trends in this experiment may spark new studies in the future.

**CONCLUSION**

The focus of the first iteration of the Aiming Game was to test the effectiveness of the different features which an emotion regulation training tool requires. In this respect, the iteration has been a success and has generated much data, feedback and comments that have been organized into a backlog on which the second development iteration is based. Play Testing with students and colleagues, Heuristic Evaluation and pilot studies carried out with the help of the Aiming Game, testing its potential as a learning- and training tool, have all led to a deeper understanding of how the development should proceed.

When working with design and collaborating on an international level, it is important to have a clear structure which allows all partners to have frequent access to the development process and state of progress. In the case of the development of this game prototype, this issue was solved by setting check-up meetings with the product owner, where initial requirements were matched.
against design choices and implementations. Having frequent check-up meetings has proven to be very beneficial for this type of development process.

When developing in a similar manner to that described in this paper, we recommend splitting large game development cycles into tangible, manageable iterations where each iteration is Heuristically evaluated by a few experts. This will, as shown in this paper, allow developers to spot design flaws early on before much time, effort and money has been spent. Play Testing is also strongly recommended but requires a product that is somewhat closer to completion since it involves game experience that can be hard to induce in early stages.

The Aiming Game will be used in several xDELIA studies regarding Emotion Regulation, in the summer of 2011. It will also be featured in a study and a paper regarding the training of emotion regulation in relation to performance in games.

ACKNOWLEDGEMENTS
This research was funded by the European Commission under the 7th Framework Programme: Digital libraries and technology-enhanced learning (ICT-2007.4.3) (ICT-2007.4.3), Project xDELIA: Excellence in public and professional decision making: boosting deliberate practice and handling biases through immersive cognitive and emotional reinforcement strategies and tools (Contract: 231830). We thankfully acknowledge the help of all our evaluation and testing participants.

REFERENCES


Desurvire, H., Caplan, M., Toth, J. A. "Using heuristics to evaluate the playability of games", CHI '04 extended abstracts on Human factors in computing systems, April 24-29, 2004, Vienna, Austria


Sennersten C. "Gameplay (3D Game Engine + Ray Tracing = Visual Attention through Eye Tracking)", Blekinge Institute of Technology Licentiate. 2008


Paper 3
A Serious Game Using Physiological Interfaces for Emotion Regulation Training in the Context of Financial Decision-making

Petar Jercic, Philipp J. Astor, Marc T.P. Adam, Olle Hilborn.

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.

Introduction

Serious games are (digital) games used for purposes other than mere entertainment (Susi, 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 (ELSPA, 2006; Mitchell & Savill-Smith, 2004; Corti, 2006; Squire & Jenkins, 2003; van Eck, 2006; see also 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 (e.g., 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, 2011) and “Darwin: Survival of the Fittest” where the player is thought options trading in the trading pit (Michael & Chen, 2006, p.151); or simulating business and stock market trading, such as a computerbased 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, 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 (cf. Rasmusen, 2007). 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 decisionmaking (e.g. Loewenstein, 2000; Adam et al. 2011a). 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 (1998) 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 decisionmaking (cf. Fenton-O’Creevy et al., 2010).

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.

**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., 2011b). 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 et al. (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 (cf. 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 influences by their emotions. Fenton-O’Creevy et al. (2010) 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 (e.g. Anttonen and Surakka, 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-O’Creevy et al. (2010) 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 1: Emotions in the valence-arousal space](image-url)
Emotion regulation can be described by the process model of Gross (1998) 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 et al. (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.

**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 (2011) gives an up-to-date overview on (serious) games using biofeedback, where we can see various usage ranging from educational support (Conati et al., 2003; Kato, 2010) to stress/relaxation/concentration training (Dupire et al., 2011; 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, 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, 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 decisionmaking 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\(^3\) 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

\(^3\) [http://www.movisens.com](http://www.movisens.com)
the xAffect framework. The arousal levels are computed in relation to a baseline period which is recorded before the game starts.

**Game description**

**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.

**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.

![Figure 2. Example sketch of game rounds](image)

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.

**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.

![Figure 3. Screenshot from the Auction Game](image)

**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.
  - 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>Level</td>
<td>Event</td>
</tr>
<tr>
<td>-------</td>
<td>-------</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 1. Game level elements varying cloud estimations**

**Evaluation**

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

**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 demonstrated. 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 & 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.

**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 (AOI) 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 et al., 2003) in order to identify suppression and reappraisal tendencies of individuals; a modified System Usability Scale (Nacke, 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.

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 (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= -.1369, SD=3.3). Other groups 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.

**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 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.

References


Technologies (ICALT), 2011 11th IEEE International Conference on (pp. 221-223). doi:10.1109/ICALT.2011.71


doi:10.1109/CW.2010.56

Paper 4
A study of the effect of emotion regulation in a shooting game using bio-feedback from electroencephalography for training in emotion regulation
Olle Hilborn, Henrik Cederholm, Jeanette Eriksson, Petar Jercic, Craig Lindley, Jennifer Petersson, Charlotte Sennersten

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 component arousal. The effect of emotion regulation tendencies on game performance was studied, and expressive suppression was found to be a significant correlational factor. When player experience was controlled for 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 by experienced players using this game, but that expressive suppression would be the most effective strategy.

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 & Shumway, 2003).

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 & Fine, 1983). 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, Vesin, Ebrahimi, & Diserens, 2008; Wolpaw, Birbaumer, McFarland, Pfurtscheller, & Vaughan, 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. recommended the use of explicit feedback since that increases immersion, positive affect, and players are able to manipulate their own physiological state. (Zander & 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.

(Kalyn, Mandryk, & Nacke, 2011) implemented a game that used different forms of biophysiological 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 (www.emotiv.com), a 16-channel EEG device specifically developed in order to be a consumer-friendly device.
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: 1. by the pattern of stimuli it emits; 2. as the behaviour required in order to perform successfully on the task; 3. as descriptions of the typical task behaviour that are exhibited during task performance; 4. as skills required to perform tasks (Wood, 1986). Throughout this article a task is described as 2, the behaviour required in order to achieve a certain level of performance. (Beal, Weiss, Barros, & MacDermid, 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% of their cognitive resources to that task, given that no ceiling effects exist for the task. In order to focus all attention towards a specific task, self-regulatory processes need to be applied. Regulation is influenced by two broad categories of effects: on-task attential 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. All of these factors, self-regulatory resources, on-task interest, and off-task distracters, vary among individuals since people will find different things interesting or distracting, as well as possess different self-regulatory capabilities.

Since emotions can be a part of off-task attentional distractions, it is important how distraction is handled. (Gross, 2002) presents a process model for emotion regulation (see Fig. 2) where two broad categories of emotion regulation are seen: antecedent-focused and response-focused. These categories are further broken down into:

Situation selection, i.e. putting oneself in situations that are comfortable and non-upsetting, e.g. a person who is suffering from gephyrophobia (fear of 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 gephyrophobic 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” (James J. 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 (J J Gross, 1998; Richards & Gross, 2000).
Wallace, Edwards, Shull, & Finch (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 (James J. Gross, 1998).

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 (J J Gross, 1998) and (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 & Timperley, 2007). In order to make this a priority within the game, the players’ emotional state directly influences the difficulty of the game. (Lindley & 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.

**Method**

**EEG**

The Emotiv EPOC Neuroheadset ([www.emotiv.com](http://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.
The Game

The game used in this study, The Aiming Game (see fig 2) (Cederholm, Hilborn, Lindley, & Sennersten, 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 (i.e. not teleporting it), 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 $30 \times \sqrt{(\text{arousal} - 1) \times R}$ 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.
Fig 2. The second level of the Aiming Game, during minimum arousal.

Fig 3. The second level of the Aiming Game, during maximum arousal.
Fig 4. The outer bounds of where the cursor was misplaced to during maximum arousal.

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.

Questionnaire
The Emotion Regulation Questionnaire (ERQ) (James J. Gross & 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 ERQ, demographic data was collected regarding age and gender. The factors “shooting game experience” and “biofeedback 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.
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.

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 1. Pearson’s correlations between experience, ERQ results and accuracy for all participants (N=48). * Correlation is significant at the 0.05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
<th>6.</th>
<th>7.</th>
<th>8.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shooting Games</td>
<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>Reappraise</td>
<td></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>Suppress</td>
<td></td>
<td></td>
<td>1</td>
<td>.361*</td>
<td>.459**</td>
<td>.483**</td>
<td>.434**</td>
<td>.469**</td>
</tr>
<tr>
<td>Accuracy Arousal</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>.759**</td>
<td>.658**</td>
<td>.564**</td>
<td>.524**</td>
</tr>
<tr>
<td>Arousal 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>.942**</td>
<td>.780**</td>
<td>.577**</td>
</tr>
<tr>
<td>Arousal 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>.813**</td>
<td>.573**</td>
</tr>
<tr>
<td>Arousal 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>.415**</td>
</tr>
<tr>
<td>Arousal 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Arousal 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 defined 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 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 2. Pearson’s correlational data between shooting game experience, ERQ score, inversed arousal accuracy, and game score for each phase. * Correlation is significant at the 0.05 level (2-tailed). ** Correlation is significant at the 0.01 level (2-tailed).

<table>
<thead>
<tr>
<th></th>
<th>Beginner</th>
<th>Expert</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exp SG</td>
<td>Reappraise</td>
</tr>
<tr>
<td>Exp SG</td>
<td>.1</td>
<td>-.153</td>
</tr>
<tr>
<td>Reappraisal</td>
<td>.256</td>
<td>1</td>
</tr>
<tr>
<td>Suppress</td>
<td>.621**</td>
<td>.144</td>
</tr>
<tr>
<td>Accuracy 1</td>
<td>.382</td>
<td>.157</td>
</tr>
<tr>
<td>Accuracy 2</td>
<td>.568**</td>
<td>.265</td>
</tr>
<tr>
<td>Accuracy 3</td>
<td>.594**</td>
<td>.363</td>
</tr>
<tr>
<td>Accuracy 4</td>
<td>.632**</td>
<td>.410*</td>
</tr>
<tr>
<td>Accuracy 5</td>
<td>.553**</td>
<td>.444*</td>
</tr>
<tr>
<td>Inv Accuracy 1</td>
<td>.465</td>
<td>.616*</td>
</tr>
<tr>
<td>Inv Accuracy 2</td>
<td>.057</td>
<td>.386</td>
</tr>
<tr>
<td>Inv Accuracy 3</td>
<td>.413</td>
<td>.635*</td>
</tr>
<tr>
<td>Inv Accuracy 4</td>
<td>.164</td>
<td>.215</td>
</tr>
<tr>
<td>Inv Accuracy 5</td>
<td>.688**</td>
<td>.596**</td>
</tr>
<tr>
<td>Phase 1 Score</td>
<td>.485*</td>
<td>.314</td>
</tr>
<tr>
<td>Phase 2 Score</td>
<td>.558**</td>
<td>.215</td>
</tr>
<tr>
<td>Phase 3 Score</td>
<td>.393</td>
<td>.308</td>
</tr>
<tr>
<td>Phase 4 Score</td>
<td>.262</td>
<td>.518*</td>
</tr>
<tr>
<td>Tot Score</td>
<td>.497*</td>
<td>.303</td>
</tr>
</tbody>
</table>

Table 3. Mean data for experts and beginners, as well as t-test results between the groups.
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 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 (J J Gross, 1998), 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.

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.

References


Paper 5
A biofeedback game for training arousal regulation during a stressful task: the Space Investor
Olle Hilborn, Henrik Cederholm, Jeanette Eriksson, Craig Lindley.

Abstract. Emotion regulation is a topic that has considerable impact in our everyday lives, among others emotional biases that affect our decision making. A serious game that was built in order to be able to train emotion regulation is presented and evaluated here. The evaluation consisted of usability testing and then an experiment that targeted the difficulty of the game. The results suggested adequate usability and a difficulty that requires the player to engage in managing their emotion in order to have a winning strategy.

Introduction
This work was a part of the Europeans project xDELIA (Xcellence in Decision-making through Enhanced Learning in Immersive Applications, www.xdelia.org), which tried to improve decision making by training emotion regulation, thus lowering impact from emotional biases. The game presented here, Space Investor, is a redesign of a simple shooting game that was affected by the players’ arousal level, called the Aiming Game (Cederholm et al., 2011). The Space Investor game was created in order to support emotion regulation training, using the concepts presented here. The other theme of this work was to evaluate if Space Investor was difficult in a way that actually could support training of emotion regulation capabilities.

Biofeedback is the concept of displaying specific bodily signals to the subject in order to make them explicit, so that the subject may change his or her behaviour. This has been a successful approach for teaching a variety of concepts such as e.g. pain regulation (deCharms et al, 2005), lowering of blood pressure (Goldstein, Ross, & Brady, 1977), and anger management (Achmon et al., 1989). Recently, studies involving games with biofeedback has been carried out where the main goal was to increase the quality of interaction by modifying the game in real time (Dekker et al., 2007; Rani, Sarkar, & Liu, 2005). These two concepts came together in the work described here by facilitating learning with a game that will encourage, by changing difficulty, a specific bio-physiological state.

Emotions can generally be classified by the independent components arousal and valence (Russel, 1980), where arousal represents excitement level and valence defines whether the arousal is positive or negative. This means that emotions can be visualized where arousal and valence defines each axis, and both need to be controlled in order to control emotions. There are methods for extracting and interpreting valence from e.g. electromyography (EMG) measuring
In order to make Space Investor a convenient, usable, and affordable game, heart rate was chosen for Space Investor to measure arousal and alter the gameplay. It is a signal that is easily measured, due to the large electric activity from the heart. Also, the sensors are very unlikely to fall off or be affected by movement since various heart rate monitors are designed on the same principles (waistband and wireless connection) that are used in sports (e.g. Gamelin, Berthoin, & Bosquet, 2006). The reason to include only arousal was to avoid an unnecessarily complex experience for the players, i.e. a multi-factored affective game system. The sensor system used is the Movisens and xAffect system (Schaaff et al., 2012), which calculates arousal and feed it into the game.

**Game Description**

The aim of the game is to navigate the spaceship from one planet to another in space, collecting resources and destroying obstacles on the way. The ship is carrying goods, which are to be delivered on the destination planet, together with the resources the player managed to pick up on the way there. The game is structured with different levels, where each level consists of a travel route between two planets. The game is presented in 3D and the camera is set to the front of a spaceship, which means that the player will not actually see the ship (First person view, see Fig. 1). Because Space Investor is a serious game that is developed through EU-funding, a very neutral setting was desirable. The choice of space and asteroids was deemed sufficient for this, since no hostile action can be taken against another living thing.
Gameplay
Space Investor is a single-player game in which the player tries to avoid getting hit by asteroids by shooting them down. The player cannot die in the game, only lose resources. The spaceship is constantly moving forward through space and is frequently approached by asteroids that must be shot down in order for the ship not to get hit. Occasionally the player will encounter resources which are collected automatically when they hit the ship, but are immediately destroyed if the player shoots at them. There is a third, hybrid type of asteroids (resource asteroids) that need to be shot at before they turn into resources.

Input
When playing the game, the ship is automatically moving forward in the space environment without the need for player input. The main goal is to aim and shoot at different object. This is done with the mouse buttons, left for the primary weapon and right for the secondary. The indirect (meaning: not under direct conscious control) input from the player is the heart rate that is registered and then calculated into an arousal value with the Movisens system (Schaaff et al.,
2012). This arousal value will distort the game and make it harder when the arousal is too high.

**Difficulty level**

When developing a serious game, such as the Space Investor game, it must be taken into consideration the fact that the target group may not be experienced game players or have the incentive to go through a steep learning curve. Therefore, the game should be designed to be playable by all types of people, ranging from hard core gamers to completely inexperienced players. In the case of the Space Investor game there is a delicate line where players must feel pressured and perceive the game as hard no matter how experienced they may be, at the same time as not perceiving the game to be too difficult for them to even try. When progressing between levels two factors will be affected.

- Asteroid spawn rate will be lower, meaning there will be shorter distance between asteroids spawning, resulting in more asteroids in less time.
- Various types of asteroids and resources will appear, making it harder to distinguish the ones that will give resources and the ones that need to be shot down.

**Arousal effects**

In order to train emotion regulation during the game, it is important that the game is sufficiently challenging in the aspect of emotional control, in order to elicit an emotional response. In the game these effects are explained with the biophysiological interface of the ship and when the physical state of the player is undesired, the ship’s functions are working at a suboptimal level. The higher arousal, the bigger each of the effects will be. The elements affected are

*Speed:* The ship’s speed is increasing if high arousal levels are detected. This makes the game harder since asteroids will hit the ship faster and also spawn with higher frequency (in time). The normal speed of the ship is 100 space units per second, with up to another 100 space units depending on arousal level. This means that the speed will double if the player has an arousal level on a maximum distance from the wanted value. The time it takes to reach the next planet is thus shorter when the player is aroused.

*Aiming offset:* A randomized offset from the aiming point will be present when unwanted levels of arousal occur. This is meant to make the game harder since the targets will be harder to hit in general. With the help of a double cosine function, a fractal movement pattern is created. The arousal
value determines the amount the fractal function is allowed to affect the
aiming. The functions are $X_n = X_{n-1} + \text{arousal} \times 50(3 \times \cos(0.6t) + \cos(\pi t))$ for
movement along the x-axis and $Y_n = Y_{n-1} + \text{arousal} \times 50(3 \times \cos(0.75t) + \cos(\pi t))$ for
movement along the y-axis, where $t$ is the time in milliseconds since the
last frame.

**Blur.** The blur is primarily a means to make it hard for the player to
distinguish between resources and asteroids, rendering the player to
mistakenly shoot down resources.

**Shields:** The spaceship has shields in order to protect it from asteroids;
these are weakened as a result of arousal levels and damage levels will
increase. Damage will make the player to lose resources and the camera
will shake (more damage equals longer shake time).

An overview of the difficulty and the specific settings used can be found in Table 1,
where the arousal effects for each level can be found, as well as the spawn rate for
asteroids. The spawn rate worked as when 600 space units had been travelled (for
level 0) a new asteroid was spawned.

**Table 4. The difficulty settings for each level. Spawn rate is measured in space units, resources
have a constant spawn rate of 500.**

<table>
<thead>
<tr>
<th>Level</th>
<th>Arousal effects</th>
<th>Asteroid spawn rate</th>
<th>Resource asteroid spawn rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>None</td>
<td>600</td>
<td>n/a</td>
</tr>
<tr>
<td>1</td>
<td>Speed</td>
<td>560</td>
<td>n/a</td>
</tr>
<tr>
<td>2</td>
<td>Speed, aiming</td>
<td>540</td>
<td>n/a</td>
</tr>
<tr>
<td>3</td>
<td>Speed, aiming, blur</td>
<td>520</td>
<td>n/a</td>
</tr>
<tr>
<td>4</td>
<td>Speed, aiming, blur, shields</td>
<td>500</td>
<td>520</td>
</tr>
</tbody>
</table>

**Weapons**
To enhance the support for an immersive experience and the drive for self
enhancement in the game, various weapons existed. This was also so that players’
could choose their own strategy when playing. During gameplay, the player was
equipped with two weapons (one for left mouse button, and one for the right one)
at all times. This made it a strategic choice about which guns the player should
choose and use. Each weapon has the two attributes damage and cooldown time,
which is the amount of time before the weapon can be fired again.
**Main Cannon:** The main cannon was the default weapon that all players received from the very beginning of the Space Investor game. It is designed to have a rather standardized functionality, which should be recognizable by most semi-experienced players, a low damage with a low cooldown time.

**Bomber:** The Bomber is a slow but powerful weapon that acts as an auxiliary weapon. It can destroy any obstacle in one hit but has a long cooldown period before it can be used again making it rather powerless against multiple targets.

**Testing of the game**
The game was tested in two ways, a small usability play testing with the focus on what players thought about the game and a larger that tested different hypotheses in the xDELIA project as well as data directly relevant to the game. The same game setup was used in both cases (see Table 1.), consisting of five levels with a length of each 18,000 space units (a maximum of 180 seconds play time, depending on arousal). It is worth noting that level four had resource asteroids that will damage the ship if not shot down, but will turn into resources if they are shot down.

**Usability test**
The play testing evaluation of the Space Investor game was done with five participants, all students at Blekinge Institute of Technology. All players filled in the game experience questionnaire (GEQ; IJsselsteijn et al., 2008) and the system usability scale (SUS; Brooke, 1996). They have both been used in earlier work to assess serious games (Cederholm et al., 2011). The answers in both questionnaires range from “1 – not at all” to “5 – completely”. See Fig. 2 for GEQ (scores can range from 1 to 5). The average result across all players regarding the SUS was 81.5, out of a total of 100.
Experiment testing

The usability and game experience was deemed sufficient to use the game for a full scale study. Sixty participants played the game once a week for three weeks. The participants’ arousal levels were recorded and classified according to the following: Arousal 0 (≤0.00), Arousal 25 (0.01-0.25), Arousal 50 (0.26-0.50), Arousal 75 (0.51-0.75), and Arousal 99 (0.76-0.99). The proportion of the asteroids that hit a player during each of the arousal states were calculated (see Fig. 3).
A Kruskal-Wallis test over the three different play sessions showed that people performed significantly better during level 1 to 4 the consecutive times they played. Moreover, it could be seen for at least two arousal states during each of these levels. Spearman’s correlation suggested significant negative correlations between number of play sessions and the probability of being hit by asteroids.

Discussion and Conclusion
Based on the SUS and GEQ, the game had adequate usability, but improvements could be made in regards to game experience. In this study it was deemed sufficient since the main purpose was not to create a game with the best game experience, but a game where emotion regulation could be practiced. Further work may go into either why some people like it or how it can be more inclusive.

From the proportion (see Fig. 3) of players that got hit by asteroids, it can be seen that the game got successively harder at higher levels, but most apparent that higher arousal values made the game harder. For each level that was modified by arousal (i.e. 1 to 4), it is evident that the optimal strategy was to lower arousal levels in order to perform better. This is in line with the purpose of the game; to encourage emotion regulation training.
Regarding the Kruskal-Wallis test together with Spearman’s correlation it was interpreted as the game became easier the more times it was played, meaning that the game is not too easy and there is room for improvement. It cannot be distinguished between if players get less stressed by the game or if they are better at regulating their emotions, since these variables are interdependent.

When Space Investor was designed, the purpose was that it should punish high arousal levels, which it very much does according to the gameplay data. During level one to four, people are twice as likely to be hit by an asteroid when experiencing high arousal compared with low. It is also progressively harder the longer the game is played. A next step in research would be to find out at what difficulty levels people start feeling too frustrated and give up, as well as when the optimal training is achieved and tune the game variables accordingly.

The conclusion drawn from this work was that while Space Investor might not be the game with impactful game experience, it is very usable. Also, it fulfils its purpose by having the optimal strategy being that players need to regulate their arousal levels in order to succeed. Learning effects were seen, but they cannot be attributed to either game adaption or a better emotion regulation capability.

References


ABSTRACT

Games are often used as training devices in various tasks, but proper biofeedback is more seldom used. Within an EU project it was explored how biofeedback games can target emotion regulation and be evaluated meaningfully. While many use games and biofeedback separately, here the focus was to combine them. This was explored through how the games were perceived and played while players were punished in-game, based on their physiological activity. By implementing games and study the interaction patterns in experimental settings, primarily correlational data was acquired. The results suggest that targeting cognitive constructs has to be validated for each specific game, since game strategies can influence the activation of the cognitive constructs.