COPYRIGHT AND CONFIDENTIALITY NOTICE

The work described in this document was performed as part of the xDELIA project (‘Boosting Deliberate Practice and Handling Biases through Immersive Cognitive and Emotional Reinforcement Strategies & Tools’), which is funded under contract No. 231830 of the European Community. The project is a collaboration between CIMNE (coordinating partner), Forschungszentrum Informatik, Open University, Blekinge Tekniska Högskola (Game and Media Arts Laboratory), Erasmus University Rotterdam (Erasmus Centre for Neuroeconomics), University of Bristol (Personal Finance Research Centre), and Saxo Bank A/S. The opinions, findings and conclusions expressed in this report are those of the authors alone and do not necessarily reflect those of the EC or any other organisation involved in the project.

No part of this publication may be reproduced or transmitted in any form or by any means, whether electronic, mechanical, photocopying, recording or otherwise; nor stored in any information retrieval system of any kind; nor used for tendering or manufacturing; nor communicated to any person, without the prior permission in writing of the copyright owner.

The contents of this document must be treated as confidential and protected as defined in the terms and conditions of the xDELIA Consortium Agreement.
Executive Summary

The principal objective of this document is to report on the results of studies carried out in Work Package 2 during Year 3, and to explain how this has contributed toward the development of the final learning pathway.

We first describe the findings of the Year 3 studies and elaborate on their relationship to the goals of the xDelia project and their implications for the evaluation of our approach to designing learning interventions.

Studies in Work Package 2 during the final year have had a primary focus on evaluating the xDelia learning pathway and its elements, which we designed in Year 2. We have carried out this evaluation with the participation of a total of 856 students and 370 private investors, who trade on their own account.

Taken together, our Year 3 studies provide much support for the validity of our approach to learning design, our learning pathway, and the learning elements which make it up.

- We show a significant impact of training on disposition effect as measured in real world trading behaviour.
- We have very positive user feedback suggesting our learning approach to be engaging, enjoyable and a good basis for learning.
- There is good evidence for the effects of the learning interventions in achieving proximal goals of improving emotion regulation, mindfulness and interoception.
- Our studies support the value of our sensor based games in diagnosing emotion regulation capabilities and the value of the Two Index Game in diagnosing a propensity to a disposition effect.

However, as is normal with such research we are also left with some puzzles which merit further research and in some cases suggest a need for further development of elements of our learning pathway.

- In some cases learning interventions lead to a paradoxical increase in disposition effect as measured by the Two Index Game. We have advanced some explanations for this finding but further research is needed to establish whether these are correct. This may involve gaining a better understanding of the ways in which playing behaviour evolves over time in repeated plays of the Two Index Game.
- Whilst we hypothesised the use of a reappraisal strategy to be associated with better performance in our emotion regulation games, in some cases better performance is also associated with use of a suppression strategy for emotion regulation. This is unexpected in the light of previous research on these emotion regulation strategies and merits further investigation.
- The online diagnostic questionnaire scales\(^1\) have mixed success in predicting financial behaviour and further development is needed here.

---

\(^1\) We refer here to the online diagnostic questionnaire used in the first diagnostic part of the learning pathway (see D18-2.4.2) and evaluated in Study S-M1/2. We did also use a range of other questionnaires in our evaluation studies.
Contents

1 INTRODUCTION ........................................................................................................ 7
  1.1 DOCUMENT PURPOSE AND SCOPE ................................................................. 7
  1.2 LIST OF ACRONYMS .......................................................................................... 7

2 KEY FINDINGS FROM YEARS 1 AND 2; WHAT WE ARE BUILDING ON ............. 9
  2.1.1 Links between emotion regulation and performance on financial decision-making tasks ........ 9
  2.1.2 Results on the disposition effect ..................................................................... 10
  2.1.3 Mindfulness approaches ................................................................................. 10
  2.1.4 Domain specific approaches to emotion regulation in a trading context .......... 11
  2.1.5 The Aiming Game (Now developed into the Space Investor Game) ................. 11
  2.1.6 The Auction Game ....................................................................................... 12
  2.1.7 The Two Index Game ................................................................................... 13
  2.1.8 The Two Index Game ................................................................................... 13
  2.1.9 Diary and reflective practice tools ................................................................. 14

3 YEAR 3 WP2 STUDIES .......................................................................................... 15
  3.1 OVERVIEW OF YEAR THREE STUDIES; AND IMPLICATIONS FOR DESIGN AND EVALUATION
    OF LEARNING INTERVENTIONS AND LEARNING PATHWAY ................................ 15
    3.1.1 User perceptions .......................................................................................... 16
      3.1.1.1 Diary ........................................................................................................ 16
      3.1.1.2 Sensor based games ............................................................................... 17
      3.1.1.3 Diary ........................................................................................................ 17
      3.1.1.4 Key issues for learning design ................................................................. 18
    3.1.2 Effects ........................................................................................................ 18
      3.1.2.1 Diagnostic validity ..................................................................................... 18
      3.1.2.2 Emotion regulation, mindfulness and interoception ................................ 19
      3.1.2.3 Financial behaviour .................................................................................. 20
    3.1.3 Key implications for learning design and further research ........................... 21
    3.2 SUMMARY OF STUDIES CARRIED OUT AND DEVIATIONS FROM PLAN .......... 23

4 FINDINGS FROM INDIVIDUAL YEAR 3 WP2 STUDIES ...................................... 28
  4.1.1 MF2: Second mindfulness study (EUR) .......................................................... 28
      4.1.1.1 Sample and methods ............................................................................... 28
      4.1.1.2 Findings ..................................................................................................... 29
      4.1.1.3 Implications for xDelia learning pathway design ..................................... 30
  4.1.2 TIG-1: Initial evaluation of Two Index Game (EUR) ........................................ 31
      4.1.2.1 Sample and methods ............................................................................... 31
      4.1.2.2 Findings ..................................................................................................... 31
      4.1.2.3 Implications for xDelia learning pathway design ..................................... 32
  4.1.3 S-M1 and S-M2: Investor feedback trial (OU, Saxo) ........................................ 32
      4.1.3.1 Sample and methods ............................................................................... 33
      4.1.3.2 Findings ..................................................................................................... 33
      4.1.3.3 Implications for xDelia learning pathway design ..................................... 38
  4.1.4 S-M3: Emotion regulation training by means of an arousal-based game (EUR) ...... 38
      4.1.4.1 Sample and methods ............................................................................... 38
      4.1.4.2 Findings ..................................................................................................... 39
      4.1.4.3 Implications for xDelia learning pathway design ..................................... 41
  4.1.5 S-M4: Improvement of Emotion Regulation by playing the Auction Game V2 (FZI-IPE) .... 41
      4.1.5.1 Sample and methods ............................................................................... 43
      4.1.5.2 Findings ..................................................................................................... 44
      4.1.5.3 Implications for xDelia learning pathway design ..................................... 46
  4.1.6 S-M5: Auction game as feedback tool on phasic physiological responses (FZI-IPE) ... 46
      4.1.6.1 Sample and methods ............................................................................... 47
      4.1.6.2 Findings ..................................................................................................... 47
      4.1.6.3 Implications for xDelia learning pathway design ..................................... 48
4.1.7 S-M6: Transfer of learned emotion regulation skills from Aiming Game to Auction Game DID and Two Index Game (FZI-IPE) ........................................................................................................................ 49
4.1.7.1 Sample and methods .................................................................................................................. 49
4.1.7.2 Findings ........................................................................................................................................ 50
4.1.7.3 Implications for xDelia learning pathway design ....................................................................... 53

4.1.8 S-M7: Mindfulness and aiming game with the Two Index Game. Influencing the disposition effect: how effective are xDelia’s learning intervention tools? (EUR)........................................................................................................ 53
4.1.8.1 Sample and methods .................................................................................................................. 53
4.1.8.2 Findings ........................................................................................................................................ 54
4.1.8.3 Implications for xDelia learning pathway design ....................................................................... 58

4.1.9 S-M8: Investor trials of the effect of mindfulness training, Auction Game and Space Investor Game training (OU, Saxo, EUR) ........................................................................................................ 58
4.1.9.1 Aim of study ............................................................................................................................... 59
4.1.9.2 Sample and methods .................................................................................................................. 59
4.1.9.3 Independent variables .............................................................................................................. 59
4.1.9.4 Dependent variables .................................................................................................................. 60
4.1.9.5 Sample dropout ......................................................................................................................... 62
4.1.9.6 Findings ........................................................................................................................................ 62
4.1.9.7 Implications for xDelia learning pathway design ....................................................................... 65

4.1.10 S-M9: Diary and reflection learning intervention......................................................................... 66
4.1.10.1 Sample and methods ................................................................................................................ 66
4.1.10.2 Findings ....................................................................................................................................... 66
4.1.10.3 Implications for xDelia learning pathway design ....................................................................... 67

4.1.11 S-HPa: Heuristic Evaluation and Playtesting of Auction Game ................................................. 68
4.1.11.1 Sample and methods ................................................................................................................ 69
4.1.11.2 Findings ....................................................................................................................................... 69
4.1.11.3 Implications for xDelia learning pathway design ....................................................................... 70

4.1.12 S-HPb: Heuristic Evaluation and Playtesting of Space Investor Game ....................................... 71
4.1.12.1 Sample and methods ................................................................................................................ 71
4.1.12.2 Findings ....................................................................................................................................... 71
4.1.12.3 Implications for xDelia learning pathway design ....................................................................... 72

4.1.13 S-HPc: Initial Usability and Playtesting of xHale Athlete .............................................................. 72
4.1.13.1 What we said we would do ........................................................................................................ 72
4.1.13.2 Sample and methods ................................................................................................................ 72
4.1.13.3 Findings ....................................................................................................................................... 73

5 REFERENCES ................................................................................................................................. 75
Table 3.1 – Summary of studies we proposed to carry out in Year 3 and any alterations to plans

Table 4.1 – Drop out over time

Table 4.2 – SM4 design overview, where NI treatment did not receive any arousal information, nor did arousal influence game difficulty. Therefore NI refers to the baseline treatment

Table 4.3 – Money earned and Time needed to reach a decision at different arousal levels

Figures

Figure 4-1: MF-2: Deviation of the optimal solution for people scoring high and low in neuroticism, and compared between groups (control versus mindfulness).

Figure 4-2: Study S-M1/2: Mean disposition effect by time and cohort

Figure 4-3: Study S-M1/2: Mean disposition effect by cohort and time (separating TIG cohort)

Figure 4-4: S-M3: Self-reported arousal (SAM) at increasing game difficulty levels

Figure 4-5: S-M3: Number of asteroids destroyed between experimental and control group

Figure 4-6: S-M4: Screenshots of the two treatment groups of study part I; with and without biofeedback (changing colours of clouds and arousal meter)

Figure 4-7: S-M4: Subjects average decision performance depending on their arousal state

Figure 4-8: S-M5: HR change subsequently after the result is revealed for a positive versus a negative outcome. The x-axis indicates the successive seconds

Figure 4-9: Mean changes from in week 2 and week 3 (compared to week 1) in state Reappraisal scores and state Suppression scores between subjects that received training and subjects that did not receive any

Figure 4-10: Mean changes from in week 2 and week 3 (compared to week 1) in HRV between subjects that received training and subjects that did not receive any

Figure 4-11: S-M5: HR drop subsequently after the result is revealed after a positive versus a negative outcome (left) on a time axis (in seconds). As can be seen, the experience of a loss (incorrect decision) results in a stronger initial parasympathetic reaction (and hence heart drop), compared to the event of a gain (correct decision)

Figure 4-12: S-M7: Change in state ERQ-reappraisal over time and per training group

Figure 4-13: S-M7: Mean state mindfulness over time per group

Figure 4-14: S-M7: Change in disposition effect over time per group

Figure 4-15: S-M8: Structure of study carried out with private investors

Figure 4-16: S-M8: Change in disposition effect by condition
1 Introduction

1.1 Document Purpose and Scope

The principal objective of this document is to report on the results of studies carried out in Work Package 2 during Year 3 and their implications for the project objectives. For completeness we also rehearse key findings from earlier in the project. The document is divided into the following sections:

- **Section 1 “Introduction”:** provides a description of the structure and scope of this document.
- **Section 2 “Key findings from years 1 and 2: what we are building on”** summarises findings from previous years and their implications for the project.
- **Section 3 “Overview of Year 3 WP2 studies and implications for design and evaluation of Learning Elements and Learning Pathway”** provides an integrated overview of Work Package 2 studies in the final year and discusses their implications for the design and evaluation of our learning approach and the elements of which it is composed. We also provide a table of studies carried out and any deviations from plans as set out in Year 2.
- **Section 4 “In Depth Studies Carried Out in Year 3”** provides a more detailed account of the individual studies in Work Package 2.

This deliverable is an outcome of Tasks 2.5 and 2.7a.

This document assumes an existing understanding of some ideas which are core to the xDelia project. In particular it assumes an understanding of emotion regulation strategies, heart rate variability as a measure of emotion regulation, the role of mindfulness in improving emotion regulation, the role of anticipated regret in biases in financial decision-making and the nature of the disposition effect. These core ideas are set out as part of our discussion of our learning design approach in deliverable D23-2.5 Pedagogical Framework (from Year 2) and in deliverable D18-2.4.2; which could be usefully read first.

1.2 List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI</td>
<td>Artificial Intelligence</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td>AUC</td>
<td>Auction Game</td>
</tr>
<tr>
<td>AOI</td>
<td>Area of Interest</td>
</tr>
<tr>
<td>AUC-DIAG</td>
<td>Auction Game - Diagnostic</td>
</tr>
<tr>
<td>AUC-DID</td>
<td>Auction Game - Didactic</td>
</tr>
<tr>
<td>BF</td>
<td>Biofeedback</td>
</tr>
<tr>
<td>BVS</td>
<td>Body Vigilance Scale</td>
</tr>
<tr>
<td>CBT</td>
<td>Cognitive Behavioural Therapy</td>
</tr>
<tr>
<td>CCT</td>
<td>Columbia Card Task</td>
</tr>
<tr>
<td>DE</td>
<td>Disposition Effect</td>
</tr>
<tr>
<td>DIAG</td>
<td>Diagnostic</td>
</tr>
<tr>
<td>ECG/EKG</td>
<td>Electrocardiogram</td>
</tr>
<tr>
<td>EEG/EMG</td>
<td>Electroencephalography/Electromyography</td>
</tr>
</tbody>
</table>

D9-2.3.3 – In-depth studies: results (Year 3)
The Emotiv EPOC is a high resolution, neuro-signal acquisition and processing wireless neuroheadset, which acquires EEG/EMG signals.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPOC</td>
<td>The Emotiv EPOC</td>
</tr>
</tbody>
</table>
2 Key Findings from Years 1 and 2; What We Are Building On

Results of Work Package 2 studies from previous years are reported in deliverable D9.2.3.2 “In-depth studies: specification (Year 3) and results (Year 2)”. In the interests of providing a more complete picture, we summarise the key findings from these studies and their implications for the project thematically below. For more detail the reader is referred to the previous deliverable.

2.1.1 Links between emotion regulation and performance on financial decision-making tasks

Taken together, the results from Year 1 exploratory studies, the regret studies and the trader field studies (RS-1, RS-2, TFS-1 and TFS-2) provide significant support for our approach to focusing on emotion regulation as an important route for improving financial decision-making.

Our experimental laboratory studies show a significant relationship between emotion regulation strategy and financial behaviour and outcomes. In particular, subjects using suppression strategies come to significantly worse decisions compared to benchmark. We also find, as in previous studies that regret associated with missed gains affects behaviour less strongly than regret associated with losses, an important explanation for the disposition effect (Fogel & Berry, 2006). However, we are also able to show that this asymmetry between behavioural responses to losses and missed gains is accompanied by an asymmetry in physiological response and that the size of physiological response and behavioural response is moderated by emotion regulation strategy. These results have fed into the design of the Auction Game. One mode of the Auction Game especially addresses this asymmetry. Players get feedback about their behavioural and physiological responses to losses and missed gains. This provides important process feedback to aid understanding of the reactions which underpin the disposition effect.

We also find in our trader field studies that emotion regulation effectiveness whilst trading (as measured by high frequency heart rate variability whilst trading; HF-HRV) rises with expertise (Fenton-O’Creevy, Davies, Lins, Richards, & Vohra, 2012) and declines in market conditions with high volatility (as measured by the VIX2 ‘fear factor’). This is an especially strong finding given the known decline in high frequency heart rate variability with age; indeed we see a decline in resting HRV with trader experience, consistent with this age decline effect). This finding coincides with the findings from our interviews with traders and investors that managing emotional reactions to market events and trading outcomes is important if they are to avoid bias, and that this is a skill which they develop over time (Vohra & Fenton-O’Creevy, 2011). This also coincides with the findings of a prior qualitative study (Fenton-O’Creevy, Soane, Nicholson, & Willman, 2010 ) that expert traders have more effective emotion regulation skills than novices and moderate performers. These HRV findings are particularly interesting in understanding adaptive behaviour since HF-HRV is an effective indicator of moment by moment adaptation to the environment. HF-HRV indexes the activity of the parasympathetic nervous system. Whilst the sympathetic nervous system is responsible for activating responses to threat and opportunity, it is the activity of the parasympathetic system which sculptures responses to provide an adaptive response to the environment.

---

2 The VIX is the Chicago Board Options Exchange Market Volatility Index, also known informally by traders as the ‘fear index’. It measures market expectations about future price uncertainty.
These studies also provided a more nuanced understanding of patterns of physiological response to information and trading outcomes. This is of direct use in the design of our learning games and approaches to delivering feedback.

### 2.1.2 Results on the disposition effect

The disposition effect is the tendency to hold assets which would sell at a loss for longer than assets which would sell at a gain. In colloquial terms an investor who suffers from the disposition effect cuts their wins and runs their losses. This bias arises out of the desire to avoid the emotional pain of realising a loss (Frydman, Barberis, Camerer, Bossaerts, & Range, Under Review). So long as the investor does not convert a paper loss into a realised loss they can console themselves that 'it will probably increase in value again'.

Existing approaches to characterising the disposition effect are primarily designed to demonstrate the existence of the effect in a given population rather than diagnose the extent to which individuals display this bias (Barber, Yi-Tsung, Yu-Jane, & Odean, 2007; Feng & Seasholes, 2005; Feng & Seasholes, 2008; Odean, 1998). Such approaches have also typically been constrained by only having trading data available on a daily basis. With access to moment by moment trading data (as for Saxo Bank clients and in a trading game) other approaches become possible. However, given the consequent volume of data, analysis presents data processing challenges.

We have developed a novel, theoretically driven, approach to measuring the disposition effect and approaches to real time calculation of disposition effect in a game environment and an approach to calculating this measure for individual investors trading histories (Yee & Lins, 2011a). A novel element of this approach involves using a market information based measure of time rather than chronological time in the calculation of asset holding times.

A simulation study and an empirical study (IFS-3) based on trading data from Saxo Clients and professional traders showed, first, that susceptibility to the disposition effect is associated with poorer financial performance and, second, that professional traders show markedly lower levels of disposition effect in their trading than private investors (Yee & Lins, 2011b).

This provides support for our view that a) the disposition effect has economically important effects on financial decision-making and b) that the disposition effect is potentially tractable to intervention since professional traders show reduced susceptibility.

This work has also provided an important foundation for the design of the Two Index Game. The Two Index Game is a game based on a simple trading task, which like most financial markets, elicits a disposition effect in the majority of players.

### 2.1.3 Mindfulness approaches

In developing approaches to helping investors develop their ability to regulate their emotions more effectively, we are investigating the efficacy of mindfulness approaches (Davidson et al., 2003; Kabat-Zinn et al., 1992) in improving emotion regulation. Mindfulness approaches use meditation based techniques to enable practitioners to develop the capacity to notice unhelpful thoughts and feelings and allow them to pass without being drawn into acting on them or continuing with them. Erasmus

---

3 The Two Index Game (see below)

4 In simple terms: this measure of time passes slower than clock time when there is little market activity and faster than clock time when there is high market activity.
University have been building here on established practice in a domain where there is significant evidence of success in enabling individuals to learn more effective emotion regulation strategies - cognitive behavioural therapy (CBT). The most recent approaches to CBT draw on the Buddhist tradition of mindfulness (Baer, 2003; Kabat-Zinn, 2003) and have had considerable success in using mindfulness training as a foundation for more effective regulation of emotions and behaviour (Davidson et al., 2003; Kabat-Zinn et al., 1992). Our innovation here is to translate key elements of these approaches into the domain of financial decision-making and developing approaches to online delivery of mindfulness training that can be delivered via a learning space attached to a trading platform.

Our Year 2 studies of mindfulness led out of Erasmus University (EUR) focused on two primary questions. First, can mindfulness interventions induce a mindful state and an improved emotion regulation in the context of a financial decision-making task? Second, can mindfulness inductions improve performance in a financial decision-making context?

Our Year 2 studies show that even brief interventions can successfully induce a mindful state, improve attention, increase capacity to monitor financial information, and improve financial decision-making.

2.1.4 Domain specific approaches to emotion regulation in a trading context

We conducted further analysis of data from trader and investor interviews conducted in Year 1 alongside data collected in TFS1 and data from a prior study (Fenton-O'Creevy et al., 2010). The purpose of this analysis was to examine approaches expert professional traders use in regulating their emotions whilst trading, as a benchmark for private investor behaviour. While it is clear that much of this emotion regulation fits generic strategies described by Gross and colleagues (Gross & Thompson, 2007), traders also deploy emotion regulation approaches which are specific to the trading domain. In particular, many traders follow a discipline of explicitly writing down trading strategies and of writing down reasons for changing strategy if they decide to make changes. This serves first as a commitment mechanism. It becomes harder to persuade yourself that you didn’t really intend the original strategy. It serves as a trigger for self-monitoring, and it provides opportunity for an intentional cognitive approach to emotion regulation (Vohra & Fenton-O'Creevy, 2011).

This analysis provides a basis for our work designing an online diary tool and reflective practice templates.

2.1.5 The Aiming Game (Now developed into the Space Investor Game)

The Aiming Game was designed and prototyped as part of Work Package 4 in collaborations between EUR and BTH. This game uses real time data from a physiological measure of arousal as input into game play. Participants shoot planes using gun sights controlled by a mouse button. Contrary to average shooting games, arousal levels are used as input to increase game difficulty; the sights shake with the level of perturbation linked to arousal. Thus to perform well in the game the player needs to regulate their arousal level. The game is intended to support learning to improve emotion regulation. Since management of physiological arousal and awareness of physiological state is closely related to effective emotion regulation this should support development of emotion regulation skills. The game supports development of emotion regulation skills in three ways. First, it provides an environment in which management of arousal levels can be practiced and rewarded. Second, by directing attention to the participant’s own physiological state it encourages improved interoception (awareness of internal physiological state; there is empirical evidence for a link between interoception and perception and regulation of emotion state...
Apart from heuristic evaluation and usability testing, one Work Package 2 Year 2 study carried out an early evaluation of the game with a group of investors. Interviews with participants showed them to be enthusiastic about engaging with this kind of game and very interested in approaches to improving emotion regulation whilst trading. We found no relationship between game performance and emotion regulation strategy in this initial evaluation. However, in this first trial we were unable to deploy our preferred sensor solution (this had not yet been fully developed) and instead were relying on a prototype EPOC EEG/EMG headset for measurement of arousal. Based on these and other trials the game was redesigned and developed into a new game (Space Investor) which retained the basic link between physiological signals and gameplay, but introduced a stronger game narrative and more sophisticated gameplay, but most importantly, the preferred sensor solution. Thus, Year 3 trials of the game (SM-6, SM-7) consistently used Space Investor with the more sensitive and better calibrated sensor solution (heart-rate).

2.1.6 The Auction Game

The Auction Game - Didactic5 (AUC-DID) was designed and implemented in close collaboration between FZI and BTH as part of Work Package 4. It uses physiological parameters to provide the players with a continuous feedback on their current level of arousal. The game was designed to help decision-makers improve two major aspects of emotion regulation. First, the game shall help players to improve interoception, i.e. the conscious awareness of one’s own emotional processes including single emotions and the overall emotional state. Second by rewarding effective emotion regulation and punishing poor emotion regulation, the game helps players to actively learn emotion regulation strategies. In order to achieve these goals, the game itself has to provide a challenging decision-making environment and elicit high levels of physiological arousal. At the same time, however, the narrative of the game has to be rather simple so it has to be playable in a limited amount of time and easy to learn for the players.

In the Auction Game - Didactic, the player is set in the position of a trader who has to buy and sell stocks. The game comprises several levels, each of which consists of a series of consecutive trading rounds. In each trading round of the game, the player can trade one single stock. In order to do so, the player is provided with three sequentially displayed price estimates for the real value of the stock. In the game, the three price estimates correspond to the ratings of three independent analysts. The player knows that the true value of the stock is always equal to the mean of the three price estimates. Therefore, the player’s task in each round is to calculate the mean of the three price estimates and then decide whether to buy or sell the stock at a given price. If the player chooses the wrong option, the payoff for the round is negative. If the trader does not chose one of the options within two seconds, the payoff for round is - €5.00. This is by design the worst payoff a player can achieve in a single round. Therefore, any action is better than no action. However, the trader can always decide to press the no-trade button and thereby jump to the next round without having to suffer from a negative payoff.

A core element of the game is that a players’ level of arousal is fed into the game to affect game difficulty. More specifically, the level of arousal (on a normalized scale between 1 and 5) has an influence on the variance of price estimates. In the game the variance of price estimates increases with arousal; thus, making it more difficult for the player to compute the mean of the price estimates. Moreover, the level of arousal is also continuously displayed by means clouds changing colour and an arousal meter. From level to level the player is exposed to more onscreen distractions, hence, making it more and more difficult to perform well in the game.

5 Also referred to as Auction Game – Emotion Regulation
(Further details of the game can be found in deliverables D14, D17 and D18. Further details on its pedagogic role can be found in deliverable D23). A detailed description is also available as conference paper (cf. Jercic et al., 2012).

Another version of the game, the Auction Game – Diagnostic (AUC-DIAG) allows an evaluation of subjects performance and corresponding physiological responses (measured in deceleratory heart rate response) to gains, foregone gains and losses. An emotion eliciting event such as information which we show in the game is accompanied by a subsequent drop in heart rate directly after the stimulus intake. The amplitude of this drop in heart rate can be used as a proxy of how negative the emotional stimulus is perceived. By assessing these phasic responses we aim to shed more light into the effectiveness of our emotion regulation interventions. Moreover, this approach allows matching physiological traits that could be related to individuals’ exposedness to decision biases such as the disposition effect.

The Year 2 Arousal Induction studies (IS5) directly helped to design the Auction Game. The game was also pretested heuristically and for usability in Year 2 with students at BTH. This game is a direct development out of the work done in the Year 2 regret studies (RS-1 (Astor et al., 2011) and RS-2) and has been evaluated and further developed in Year 3 studies.

2.1.7 The Two Index Game

The Auction Game was designed and implemented in close collaboration between FZI and BTH. This game uses real time data from heart rate to detect arousal. In order to maximize their profit, players engage as traders in a series of buy and sell decisions, where they are given price estimations for a good that they can buy or sell. The Auction Game is an emotion regulation training game, which has gameplay closer to the context of trading. By feeding the level of arousal back into the game improved emotion regulation is rewarded with increased accuracy of price information. From level to level the player is exposed to more distracters. It hence will become more and more difficult to perform well.

Another mode the game provides players with feedback on their individually assessed behavioural and physiological reactions to information on gains, losses and forgone gains. This provides important process feedback to aid understanding of the reactions which underpin the disposition effect.

(Further details of the game can be found in deliverables D17 and D18. Further details on its pedagogic role can be found in deliverable D23).

The Year 2 Arousal Induction studies (IS5) directly helped to design the Auction Game.

The game was also pretested heuristically and for usability in Year 2 with students at BTH. This game is a direct development out of the work done in the Year 2 regret studies (RS-1 and RS-2) and has been evaluated and further developed in Year 3 studies.

2.1.8 The Two Index Game

The Two Index Game engages players in a trading game which abstracts some key elements of the trading task. One index provides information which predicts with both lag and error the price of a second index. Players have the opportunity to buy and sell (and later short sell) units of the second index. The game includes an artificial intelligence (AI) module which plays the game in parallel using only information available to the player and making optimal use of that information. The players score is calculated as a percentage of the optimal score achieved by the AI. The game also calculates the level of disposition effect shown by the player in real time. This can be either fed back at the end of the game or displayed in real time.
The Year 2 disposition effect studies (IFS-3) fed directly into the design of the Two Index Game. In Year 2 the game was evaluated heuristically and for usability. Early results (e.g., SM-11) showed the game to elicit a disposition effect in the large majority of players and to produce a good distribution of disposition effect scores.

### 2.1.9 Diary and reflective practice tools

Our work in Year 2 on domain specific approaches to emotion regulation (Vohra & Fenton-O’Creevy, 2011) together with other engagements with stakeholders led to our understanding of the role of ‘writing things down’ as a commitment mechanism, and a support for emotion regulation, as well as a support for critical reflection.

This led to our decision to develop a trading diary approach as a support for the transfer stage of our learning pathway. This also implied the need for further work in Year 3, to develop our understanding of the optimal design for such a trading diary.
3 Year 3 WP2 Studies

3.1 Overview of Year three studies; and implications for design and evaluation of Learning Interventions and Learning Pathway

Studies in Work Package 2 during the final year have had a primary focus on evaluating the xDelia learning pathway and its elements, which we designed in Year 2. (A detailed account of the learning design may be found in the Year 2 deliverable D23-2.5: Pedagogic Framework, and in the end of project deliverable D18-2.4.2: Intervention package – specification and development).

We have carried out this evaluation with the participation of a total of 856 students and 431 private investors, who trade on their own account.

As we noted in our Description of Work and previous deliverable, D9-2.3.2, evaluating user engagement in an end to end learning pathway would not be achievable within the scope of the project.

“Given that the ideal learning pathway through this learning design is around 6 months in duration, it is not practical within the scope of the project to follow a large cohort of learners through the entire learning design (See D23 – Pedagogic Framework). Rather we primarily examine each element of the learning design through a series of studies to establish whether each element performs as expected. Where we see important synergies between different elements of the design, we test their joint effects.”

Thus our approach has been to evaluate individual elements of the learning pathway and, as possible, combinations of these elements.

We have carried out this evaluation at two main levels: evaluating user perceptions of the learning experience, and evaluating effects of learning experiences.

Data on user perceptions has been gathered through a combination of surveys and interviews with participants in trials of learning elements and has included data on:

- Usability
- User engagement and enjoyment
- User perceptions of learning outcomes and learning potential

Data on effects has where possible been gathered through the use of studies which employ a randomised control design. In particular we have examined the effect of learning interventions on:

- Improved emotion regulation
- Improved mindfulness
- Improved interoception and body awareness
- Financial decision-making behaviour (including susceptibility to disposition effect.)

Recruiting private investors into randomised control studies poses a substantial challenge. Whilst this group represent our primary target audience for the learning interventions, their profile is such that there are few affordable incentives to attract them into our studies beyond the intrinsic value participation may have for their own learning. This presents challenges for creating robust research designs which can demonstrate effects. Thus we have taken a two pronged approach to recruiting participants to studies. Initial studies (especially of sensor based games) have used student participants. This has increased our...
ability to design robust studies which use affordable incentives for participation and performance. These have then been supplemented by studies with investors. These have the benefit of greater ecological validity since they draw on our target audience rather than lay decision-makers with little or no experience of trading. However, such studies involve some design compromises (e.g. investors are unlikely to engage with studies where they are part of a control group with no meaningful learning benefits) and are more prone to participant drop out given the lack of viable external incentives.

We summarise our key findings below. A more detailed account of each study may be found in Section 3.3.

Before discussing key findings it is appropriate to set out our position regarding the role of emotions in decision-making. It is common in the judgment and decision-making literature (e.g. Bazerman, 2002; Kahneman & Tversky, 1979) to treat emotions as either a component of future utility or a disturbance to optimal decision-making. In contrast our position is that as noted in a recent major review (Phelps, 2006), emotions are inextricably intertwined with cognition at all stages of decision-making and play an important adaptive role in fast and frugal decision-making. As noted in an early paper developed as part of this project (Fenton-O'Creevy, Soane, Nicholson, & Willman, 2011),

“To ask whether emotion disturbs or aids traders’ decision making is to ask the wrong question. Traders’ emotions and cognition are inextricably linked. Therefore a more productive question to ask in this context is whether there are more or less effective strategies for managing and using emotion in financial decision-making.”

Thus our learning design does not seek to eliminate the role of emotion in financial decision-making, nor to replace heuristic decision-making with (slower more cognitively costly) normatively rational analysis. Rather, we propose that the important question for traders is not whether they have emotions, but how they engage with them. Our work in the final year of this project builds on evidence from earlier in the project, that a key facet of trader expertise is the development of their emotion regulation skills (Fenton-O’Creevy et al., 2012; Fenton-O’Creevy et al., 2011).

Our learning games primarily support development of emotion regulation capabilities through increasing self-awareness of internal states (interoception). Thus the focus is not the elimination of emotion, but greater self-awareness and increased ability to reflect critically on emotion-informed choices.

3.1.1 User perceptions

Work on evaluation of user perceptions involved collaboration across Work Package 2 and Work Package 6. The Work Package 6 “Evaluation Reports” deliverable D20-6.3.3 provides a more detailed account of the evaluation of user perceptions. However, for completeness we also summarise these findings here.

Feedback from participants on the learning value of our learning elements and their engagement with them was, overall, very positive, although often provided useful help in thinking about how to improve our approach.

3.1.1.1 Diagnosis and feedback

The diagnosis and feedback stage of the learning pathway uses an online diagnostic questionnaire and delivers personalised (multi-media) feedback to participants on their emotion regulation strategies and the extent of disposition effect they show in their trading history, alongside didactic materials on emotion regulation and trading biases. They also had the opportunity to play an online trading game (the Two Index Game), which provides game based feedback on susceptibility to the disposition effect. 270 Saxo clients took part in a trial of these learning materials (S-M1/2: section 3.3.3).
The large majority of participants were positive about the value of the learning intervention. They found the learning materials informative (76-82%), engaging (80-82%), provided insight into their own trading (85%) and a good basis for reflective learning (76-78%).

Those who took part in the online trial of the Two Index Game felt it was fun (69%) and in combination with other learning elements, on balance, likely to help them reduce their disposition effect in real world trading (58%). In a follow-up study (S-M8) 70% agreed that they would recommend it to a fellow trader.

### 3.1.1.2 Sensor based games

The primary role of the two sensor based-games (Auction Game and Space Investor) is as learning environments to practice and develop emotion regulation skills (with a subsidiary role in diagnosis of emotion regulation skill). In these games, game difficulty is linked to level of player arousal (as measured by heart rate); thus rewarding effective regulation of arousal whilst playing a fast-paced and arousing game. The most important user experience data on these games comes from study S-M8, since this study drew on participants from the target audience (private investors). However, user experience data was also collected on student participants in studies S-M3 and, S-M4 and fed back into game development. User experience data was also collected as part of the cycle of play testing and heuristic evaluation carried out as part of Work Package 4 and reported as study S-HP below.

Feedback from investor participants on the sensor games was very positive. Nearly all (80-82%) found them engaging and fun and 95% felt xDelia games could help them learn to manage their emotions more effectively. All participants agreed that their emotions affect their financial decision-making and that better emotion management could improve their trading.

There were a few exceptions to the positive feedback. Some of the younger (student) participants in play testing found the games insufficiently challenging and engaging (even once difficulty was increased in early changes). Two elderly men (experienced day traders) were very negative about our ‘video game’ approach (referring to the Space Investor and Auction Games) becoming very cross at the claim that such games could be relevant to trading. Further conversation revealed a high degree of anxiety at playing such games and a sense that their difficulty in getting to grips with the game was somewhat humiliating for them. Thus it is apparent that familiarity with and skill at video game play at both ends of the experience spectrum provides some limits on application of the approach for some audiences.

### 3.1.1.3 Diary

Our overarching learning design gives emphasis to the problem of learning transfer from the game setting to the practice of trading. We have proposed that the use of trading diaries should play a key role in transfer, by providing the opportunity to record and reflect on emotion states and emotion regulation success and by acting as a ‘commitment mechanism’ for trading strategies. Data on user views of this element were primarily collected in S-M9.

The majority of participants were positive about the value of recording emotions in the xDelia Trading Diary, with 67% agreeing that keeping the diary helped them manage their emotions effectively during trading (22% undecided, 11% disagree, 0% strongly disagree). 67% felt that recording their emotional state before and after trading made them aware of their emotions (22% undecided, 11% disagree, 0% strongly disagree) and 50% felt that this awareness of their emotions during trading helped them make better trading decisions (44% unsure, 6% disagree). Part of the role of the xDelia Trading Diary is as a reflective mechanism, and 67% of participants reviewed their previous diary entries and felt that reviewing their previous diary entries provided insights into the role of their emotions in their trading decisions.

In a follow-up focus group participants confirmed the value of a diary for reflective learning but also made apparent the challenges in following the discipline to keep a diary over time. Many particularly emphasised the value of automated data capture from a trading platform significantly reducing the effort.
needed to keep a diary and increasing the potential to ‘drill down’ to look at specific episodes when reflecting on diary entries.

### 3.1.1.4 Key issues for learning design

User feedback, both through surveys and interviews provided strong confirmation that our learning elements were engaging and promote reflective learning. Further, participants from our target audience (private investors who regularly trade on their own account) provided strong confirmation of the importance of supporting the development of emotion regulation skills.

While level of familiarity with video game does influence the acceptability of our games to participants, in their latest versions they seem well calibrated to the skills of the majority of our target audience. Further development might usefully provide for games to have configurable difficulty settings or to self-configure depending on player performance.

The diary approach clearly has potential in supporting transfer and Saxo Bank is working on incorporating features such as the automatic data capture of information about market action and trading at key points which participants indicated would be so valuable. However, implementation of such an approach must clearly be customised to the particularities of specific trading platform environments.

### 3.1.2 Effects

#### 3.1.2.1 Diagnostic validity

Three main elements of the xDelia learning pathway were intended to have a role in diagnosis and feedback. The Two Index Game provides a game based measure of disposition effect, the Online Diagnostic Questionnaire provides diagnosis of habitual emotion regulation strategies and the two sensor based games (the Space Investor and the Auction Game) can be used to provide feedback and train emotion regulation capability.

##### 3.1.2.1.1 Two Index Game

The Two Index Game was designed primarily as a game space in which traders could learn about a particular bias (the disposition effect) and practice emotion regulation skills without risk to real world trading performance. However, we also felt that it might, for traders with little existing trading record, also offer the possibility of diagnosing a propensity to a disposition effect. Early work in Year 2 had already shown the incidence of disposition effect in gameplay to show a similar distribution to disposition effects typically detected in market data.

In Year 3, feedback trials (S-M1, S-M2) offered the chance to examine the correlation between disposition effect in gameplay and in actual trading. We found a significant positive correlation between disposition effect scores calculated from participants’ real trading history and gameplay ($r = 0.28; p < .05; N=58$). Further this correlation with game based disposition effect remained consistent across three separate periods of trading history.

While the correlation is modest, there are good reasons why we should not expect a very high correlation. First, we know that disposition effect is affected by context. While the game context was standardised across participants, participants were, in their real world trading, trading a wide range of different assets and facing dissimilar and changing market conditions. Secondly, separate findings (see S-M8) suggest that it may take several plays of the Two Index Game before participants’ master game mechanics sufficiently to really focus on the trading task. In the feedback trial participants only played the game once. That we find a relationship between a game based measure (where the only incentive is a desire to get a high score) and a real world trading context with significant sums of money at stake is highly notable. We also find
the Two Index Game measure of disposition effect to have reasonable test re-test reliability (S-M6, S-M7, S-M8) (e.g. S-M8 intraclass correlation coefficient = 0.74, p<.000).

These results not only provide validation of the game as a diagnostic measure but also are an important contribution to our understanding of the disposition effect suggesting such biases to have a trait component as well as being influenced by context.

3.1.2.1.2 Sensor based games
Year 3 studies also provided support for the diagnostic validity of the sensor based games. S-M4 also shows a significant correlation between participants self-ratings of habitual emotion regulation strategies and game performance, although the data also suggest that the game rewards reappraisal and suppression strategies equally (we hypothesised reappraisal to be more effective).

S-M5 also shows emotion regulation as measured by balance in physiological responses (i.e. heart rate) to gains and losses to predict game performance.

3.1.2.1.3 Online Diagnostic Questionnaire
While the questionnaire scales used in the diagnostic survey are well tested and well understood and the used of this questionnaire was part of an apparently successful learning intervention (as measured by change in disposition effect), the results on relationship between questionnaire measures and trading behaviour were disappointing. In S-M1/2 there was no significant correlation between questionnaire measures and disposition effect. Although measures of emotion regulation (in the ERQ) did show a significant correlation with holding times for both assets in gain and assets in loss. S-M6 also showed mixed results. Other studies (TIG-1, S-M8) did show a relationship between effective emotion regulation and disposition effect as measured in the TIG. On balance, the best we can say is that these results on the questionnaire are inconclusive.

We consider there to be two plausible explanations of this result. A) It is possible that investors were more likely to answer positively to questions about emotion regulation if they experience a significant need to regulate emotions during their trading. In other words the ERQ may be picking up both approach to emotion regulation and amount of regulation. B) Reappraisal may be associated with different goals. If the goal is to manage emotional reactions in the interest of a balanced approach to decision-making then we would expect a reduction in disposition effect. However if the goal is to avoid negative emotions by avoiding responsibility for outcomes then it could lead to greater disposition effect. It is also possible that emotion regulation has no relationship to disposition effect. However, given results in other studies which do find a relationship (including a study conducted in parallel to xDelia at OUBS) it seems likely that there are contingencies which our trail has not accounted for.

These results suggest that the online diagnostic questionnaire should be developed further to improve validity in relation to trader behaviours. First, we suggest recasting the ERQ questions to focus on habitual approaches to emotion regulation in everyday life. Second, we suggest developing a dedicated scale which measures traders own perceptions of susceptibility to emotion effects on their trading. (This work is now ongoing).

3.1.2.2 Emotion regulation, mindfulness and interoception

3.1.2.2.1 Sensor based games
A first and basic concern in evaluating our learning games was to establish whether they behave as designed and provide a valid learning environment.

Studies S-M3 and S-M4 addressed this question for the Space Investor and Auction Games respectively. In both games better emotion regulators got higher scores, participants self-reports of arousal correlated with game based measures of arousal, and participants reported that paying attention to their arousal levels helped them play the games better. Game performance also, as intended, is higher for those who
showed greater success in controlling their levels of physiological arousal. Study S-M4 also showed emotion regulation (as measured by high frequency heart rate variability) to predict management of arousal levels and hence performance in the game.

Study S-M3, did though reveal uneven levels of difficulty across game levels rather than a steady increase in difficulty as intended. The study also raised issues about the use of discrete bands of arousal levels rather than a continuous arousal measure in the Space Investor game. This led to revisions to the game design and a slight change in Space Investor in level difficulty and operationalisation of arousal values to feed into the game for the SI-version that was eventually used in studies S-M6 and S-M7.

A second question was whether these games would be successful in supporting the improvement of emotion regulation skills. Evidence on this question comes from studies S-M6 and S-M7 (with students) and S-M8 (with private investors).

S-M6 finds that effectiveness of emotion regulation as measured both by questionnaire and heart-rate variability (AUC-ER) does improve over a three week training course using sensor based games.

S-M7 was similar to S-M6 but used the Space Investor game instead. Moreover, in this study, a mindfulness exercise was introduced into the course. As hypothesized, playing the Space Investor repeatedly over a three-week period was associated with increases in self-reported interoception behaviours (i.e., attention to bodily cues). Although trait levels of emotion regulation remained unaffected by Space Investor or mindfulness, state emotion regulation (RA) did demonstrated increased levels compared to controls for people who played SI. Only the mindfulness group showed a significant and consistent decline throughout the three trials. As mindfulness is well-associated with improvements in emotion regulation, this likely reflects a shift of attention by participants from RA-strategies to increasing levels of interoception (i.e., via the mindfulness exercise that involved a strong focus on bodily cues). The finding that body vigilance (i.e., attention to bodily signals) improved consistently in all conditions compared to controls, supports the view that the current approaches were effective in improving interoception, which is an important prerequisite for emotion regulation.

S-M8 provided further evidence for the impact of the sensor games, showing gameplay to be associated with improvements in interoception (although these improvements seemed independent of the use of biofeedback in the games).

Our Year 3 studies also sought to integrate evaluation of the role of mindfulness training. S-M7 shows our training interventions to improve state mindfulness and interoception (in a student population) and S-M8 provides support for the validity of the training approach with a private investor population.

3.1.2.3 Financial behaviour

Finally, we also sought to evaluate whether our training approaches bring about change in financial behaviour. In most cases we used the Two Index Game as our measure of financial behaviour (access to individuals’ real trading data is very difficult and such data does not exist for student participants). However, in one study (S-M1/2) we did have the opportunity to take a before and after measure in relation to real trading behaviour.

In our feedback trial study (S-M1/2), which involved diagnosis and feedback on habitual emotion regulation strategies, and on disposition effect with accompanying learning materials we measured disposition effect in actual trading data for each participant at three time periods; before the training and in two periods after. We found no significant difference between the training group and a comparison group prior to the training but significantly lower disposition effect in the training group (and relative to the comparison group) following the training. This represents fairly strong evidence for the impact of our training approach on the disposition effect. Given the failure of prior ‘debiasing’ research to show such effects, especially outside the laboratory (Bazerman, 2002; Fischhoff, 1982; Lilienfeld, Ammirati, & Landfield, 2009), this is an important vindication of our approach.
Intriguingly and perhaps importantly for banks who deploy our approaches, there is evidence that clients who participated in our training intervention may have persisted as actively trading bank clients for longer than a comparison group.

Results on the impact of training on the disposition effect as measured in the Two Index Game are less positive. While S-M8 provides some evidence for a post training fall in disposition effect, both S-M5 and S-M7 show increases in disposition effect after training (and S-M8 shows an initial rise in DE).

One explanation is that self-regulation capability is like a muscle and can become exhausted (Muraven & Baumeister). Since both the sensor games and the Two Index Game require active self-regulation, it is possible that the short term effect of playing the sensor games (which require active self-regulation) is to exhaust self-regulation capacity, whilst learning effects take time to build. Participants playing the Two Index Game immediately after the sensor games then have depleted self-regulatory capabilities leaving them more prone to emotion driven biases.

We do though see an increase in disposition effect across multiple plays of the Two Index Game in several studies. One possible explanation is suggested by comments made by several participants in follow-up interviews. A typical comment was “I wasn’t really being myself as a trader until I had played it a few times”. Further questioning suggested that in initial plays they were quite focussed on gameplay mechanics. It was only having mastered these that they felt free to focus on making as much money as possible. Disposition effect is understood to be founded in emotional reactions to gains and losses; and thus most likely to occur when the player is focused on levels of gain and loss rather than on game mechanics. Thus, it may be that the diagnostic value of the game is strongest after sufficient play to learn the game mechanics.

### 3.1.3 Key implications for learning design and further research

Taken together, our Year 3 studies provide much support for the validity of the learning approach and the learning elements which make it up.

- We have very positive user feedback suggesting our learning approach to be engaging, enjoyable and a good basis for learning.
- There is good evidence for the effects of the learning interventions in achieving proximal goals of improving emotion regulation, mindfulness and interoception.
- Our studies support the value of our sensor based games in diagnosing emotion regulation capabilities and the value of the Two Index Game in diagnosing a propensity to a disposition effect.
- We show a significant impact of the initial stage of our learning pathway on disposition effect as measured in real world trading behaviour.

However, as is normal with such research we are also left with some puzzles which merit further research and in some cases suggest a need for further development of elements of our learning pathway.

- In some cases learning interventions lead to a paradoxical increase in disposition effect as measured by the Two Index Game. We have advanced some explanations for this finding but further research is needed to establish whether these are correct. This may involve gaining a better understanding of the ways in which playing behaviour evolves over time in repeated plays of the Two Index Game.
- Whilst we hypothesised the use of a reappraisal strategy to be associated with better performance in our emotion regulation games, in some cases better performance is also associated with use of
a suppression strategy for emotion regulation. This is unexpected in the light of previous research on these emotion regulation strategies and merits further investigation.

- The diagnostic questionnaire scales have mixed success in predicting financial behaviour and further development is needed here.
# 3.2 Summary of studies carried out and deviations from plan

The following table summarizes the studies we proposed to carry out in Year 3 and any alterations to plans. The first five columns provide the study specification which appeared in deliverable D9-2.3.2 setting out the objectives of Year 3 for Work Package 2. The final column describes any deviations from those plans with reasons for the deviation.

<table>
<thead>
<tr>
<th>Id</th>
<th>Study title</th>
<th>Objectives and use</th>
<th>Starts - Ends</th>
<th>Lead partner (Collaborators)</th>
<th>Deviations from plan</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-M1</td>
<td>Investor feedback trial A</td>
<td>Study evaluates views on providing feedback to Saxo clients on their propensity to a disposition effect, alongside data on emotion regulation and mindfulness. Evaluation data will be collected on investor views of the value of such feedback and interest in learning opportunities which build on the feedback.) The trial also provides an opportunity to validate the feedback measures in relation to actual trading behaviour by examining the relative strength of the relationship of these different measures to disposition effect among Saxo clients (as measured by actual trading) and others (as measured by index game) to compare their utility as feedback instruments.</td>
<td>Y3Q1 - Y3Q2</td>
<td>OUBS, SAXO (OU-IET)</td>
<td>The major deviation involved combining the two studies and running them together so that we could (as requested by reviewers) run multiple elements of the learning pathway together. In the process of implementing these studies, the strategic profile of the xDelia project was raised significantly in Saxo Bank. This had a very positive effect on opportunities for exploitation but meant further internal hurdles to sign off study elements. In consequence some elements of data collection for these studies ran on into the final quarter and collecting data on actual trading disposition effect subsequent to the feedback trial was completed in the last month of the project in order to maximise the follow-up trading period.</td>
</tr>
<tr>
<td>S-M2</td>
<td>Investor feedback trial B</td>
<td>Online delivery of index game to Saxo clients. a) Evaluation of the game as a didactic approach to learning about the disposition effect. Evaluation data will also be collected on investor experience of such feedback and game based learning opportunities which build on the feedback.) b) The trial also provides an opportunity to validate the game as a diagnostic instrument in relation to actual trading behaviour. A follow-up will re-examine participants' disposition effect in their trading data to examine effects of the intervention relative to a control group.</td>
<td>Y3Q2 - Y3Q3</td>
<td>OUBS (SAXO, OU-IET)</td>
<td></td>
</tr>
<tr>
<td>S-M3</td>
<td>Emotion regulation training by means of an arousal-based aiming game</td>
<td>The aiming game uses an established measure of arousal (heart rate) as input to influence gameplay performance. The major aim of this experiment is to verify whether emotion regulation can be trained using the aiming game. The hypotheses are that by playing the aiming game extensively, emotion regulation can be trained, and game performance will improve. Improvements in emotion regulation will be examined both in terms of success in managing arousal and heart rate variability. The study will be conducted with students at the labs of EUR, and provide input for the learning intervention to determine how the aiming game is best used within the intervention package.</td>
<td>Y3Q2 - Y3Q2</td>
<td>EUR (FZI-ESS)</td>
<td>This study used the new version of the Aiming Game called the Space Investor game (SI). This used a different physiological input signal (ECG) from the previously used EPOC (EEG) signal and was situated in a more neutral game context (players target asteroids rather than planes). To reflect this change the word “aiming” was removed from the study title. Otherwise the study was carried out as planned.</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>S-M4</td>
<td>Improvement of Emotion Regulation by playing the Auction Game V2</td>
<td>The Auction Game V2 uses an established measure of arousal (heart rate) as input to influence gameplay performance. The goal is to improve emotion regulation by playing the game for an extended period. Relevant questions for this study are: Do players improve their performance in the game and their skills in emotion regulation in a game that shows similarities to financial trading? Do players perform better in a subsequent task that checks for propensity to the disposition effect (Two Index Game)? Improvements in emotion regulation will be examined both in terms of success in game performance and heart rate variability. In a subsequent task participants’ propensity to the disposition effect (with the Two Index Game) is compared to a group that did not play the game intensively? The study will be conducted with students at the labs of FZI, and provide input for the learning intervention to determine how the Auction Game V2 is best used within the learning intervention package.</td>
<td>Y3Q2 - Y3Q2</td>
<td>FZI-IPE (FZI-ESS/BTH/SAXO)</td>
<td>This study remained focused on the differences in HR, HRV, AUC performance and emotion regulation strategies. This was not extended to include the evaluation of the TIG. This ensured a clearer focus on the main objective of the study. The examination of the influence of learning interventions on the TIG was shifted to S-M5 and S-M6.</td>
</tr>
<tr>
<td>S-M5</td>
<td>Auction game as feedback tool on phasic physiological responses</td>
<td>Evaluate Auction Game V 1 as a learning tool in relation to emotional overweighting of losses. The goal is to improve emotion regulation by playing the game for an extended period. Game players will learn by playing this game that their financial decisions have a significant input towards emotions and vice versa. Triggered</td>
<td>Y3Q2 - Y3Q3</td>
<td>FZI-IPE (FZI-ESS/BTH/SAXO)</td>
<td>Study was carried out as planned.</td>
</tr>
<tr>
<td>Project Number 231830</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Becoming aware of emotions in financial decision-making and learn to regulate these emotions by playing Auction Game V1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotions will be joy/pride, frustration, pride and regret. Subsequently the players get feedback on behavioural and physiological responses and their implications for performance in financial decision-making. The main questions here are (1) can people learn to deal with their emotional phasic responses and (2) does this group perform better in a subsequent task relating to the disposition effect (measured by the Two Index Game) compared to a group that did not play the game intensively? The study will be conducted with students at the labs of FZI, and provide input for the learning intervention to determine how the Auction Game V2 is best used within the intervention package.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| S-M6 |
| Mindfulness, aiming game, auction game, and index game – combined study. Transfer of learned emotion regulation skills from Aiming Game to Auction Game V2 and Index game |
| Can the learned skills of emotion regulation be transferred from one task to another? The Aiming Game is more fundamental, where the Auction Game already is more related to real financial decisions. Questions to be answered by this study are: (1) Can players which improved their emotion regulation by playing the aiming game transfer these emotion regulation skills towards auction game V2 compared to a control group (a control group is already available from S-M5) |
| Y3Q3 - Y3Q3 |
| FZI-IPE (EUR, OUBS) |
| This study used the new version of the Aiming Game called the Space Investor game (SI). This study did not deviate from its original objective. Due to time constraints, it made sense to first concentrate on the Auction Game thereby complimenting SM4. Findings were also strengthened by including the use of the emotional stroop task to assess participants information-processing and the Weber task as an additional measure of participant’s disposition effect. |

| S-M7 |
| Mindfulness and aiming game with index game. Influencing the disposition effect: how |
| Two elements for the learning intervention within xDelia are a) emotion regulation using biofeedback (aiming game/auction game), b) mindfulness interventions. In a series of studies in students, EUR will investigate the effectiveness of both interventions in students, and will test whether using a combined approach of both mindfulness techniques and emotion regulation training is the |
| Y3Q2 - Y3Q3 |
| EUR (FZI-ESS) |
| The original aim of the study remained the same but the major deviation was that EUR collaborated with FZI to complement findings from S-M6. This increased the statistical power by being able to adopt a longitudinal approach to examining the effects of emotion regulation |

D9.2.3.3 – In-depth studies: results (Year 3)

© xDelia Consortium – All Rights Reserved
<table>
<thead>
<tr>
<th>Project Number 231830</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S-M8</strong> Investor trials of the effect of mindfulness training, Auction Game and Aiming Game training depending on outcomes of student trials above**</td>
</tr>
<tr>
<td>Dependent on the outcome of the student studies, an evaluation will be implemented with a cohort of investors. The main goal is to examine whether findings from the lab settings can be extended to xDela’s target population (i.e., investors) and determine whether findings on the most effective elements from the mindfulness/aiming game approach to influence the disposition effect, can be replicated in investors. A follow-up will re-examine participants’ disposition effect in their trading data to examine effects of the intervention relative to a control group.</td>
</tr>
<tr>
<td><strong>Y3Q2</strong> EUR (OUBS, SAXO)</td>
</tr>
<tr>
<td><strong>Y3Q4</strong></td>
</tr>
<tr>
<td>This study used the new version of the Aiming Game called the Space Investor game. To reflect this change the word “aiming” was replaced with “space investor” in the study title. The study was also split into 3 different studies that all contributed toward evaluating the outcomes from the student studies with our target population (i.e. private investors). Two pilot studies were carried out at the FOREX and the Trader Expo investor trade show. These events were also used to recruit investors to take part in a 3 week longitudinal study.</td>
</tr>
<tr>
<td><strong>S-M9</strong> Diary and reflection LI including capture of HRV data whilst trading**</td>
</tr>
<tr>
<td>A cohort of investors engage in a diary and physiological monitoring based reflection process (over approx. 6 weeks). The learning intervention is aimed at skill development &amp; at learning transfer. We will capture data on investor experience and nature of diary entries.</td>
</tr>
<tr>
<td><strong>Y3Q1</strong> OU-IET (SAXO)</td>
</tr>
<tr>
<td>A paper and pen prototype of the diary tool was developed and trailed during SM-8c for a period of two weeks. This enabled testing of the three key purposes indicated in D18_2.4.1, Section 3.1.8. Further this prototype enabled the testing of the following key features of the planned online diary: i) Reflection on trading practice via time stamped entries, ii) Recording of emotional state before, during and after trading, and iii) Recording compliance of and reflection upon mindfulness training and its training on the disposition effect; enabled learning effects to be examined over a three weeks period of several emotion regulation strategies (e.g., biofeedback game Space Investor, mindfulness exercise) and examined their disposition effect (TIG) over time.</td>
</tr>
</tbody>
</table>
Heuristic evaluation will continue in parallel to other studies as study results feedback into further iterations of development. The heuristic evaluation focuses on usability and game play and is performed early in each development cycle to guide the development. This kind of evaluation is an expert evaluation performed by game design and usability specialists. In addition each development cycle ends with a play testing study where the focus of the evaluation is from the perspective of the player, i.e. how the player experiences the game and the gameplay.

Studies carried out as planned, primarily reported in deliverable D14.

Table 3.1 – Summary of studies we proposed to carry out in Year 3 and any alterations to plans
4 Findings from individual Year 3 WP2 studies

Section 3.1 provided an integrated overview of the findings from Year 3 Work Package 2 studies and their implications. This section provides further detail by summarising the main results for each individual study. Given the number of studies, we have provided an abbreviated account of each study in this deliverable. We have tried to provide sufficient detail to cover the scope of work carried out and validity of results without making this document unreasonably long. Full accounts of studies will be produced as conference and journal articles. Where these exist already they can be accessed via the project website www.xdelia.org.

4.1.1 MF2: Second mindfulness study (EUR)

This study was a Year 2 study (additional to original plans) which was not complete in time for our deliverable (D9-2.3.2) with the other Year 2 studies. We report it here for completeness.

Our previous work (MF1) showed that mindfulness interventions successfully increased state mindfulness within Business Administration students who may hold an equally cynical view of meditation approaches compared to xDelia’s target population (i.e., private investors). More interestingly, we observed that a short exercise in mindfulness influenced financial decision-making by making them more sensitive to financial information on gains and losses. One drawback of our earlier work was that student participation was voluntary, and they received no financial compensation based on their performance on the task. Hence, study results may vary when people are making actual financial decisions in the sense that there is something at stake. Second, it could be argued that decision-making in the tasks presented in MF1 are not comparable to real world financial decision-making in the sense that there is room for cognitive deliberation on the outcomes of each decision. This could significantly reduce the impact of emotions in decision-making and allow for more rational decision-making, which in turn could temper associations between emotion regulation exercises and decision-making as no emotions were present to begin with.

In this study, we extend the findings of MF1. Remuneration largely depended on game performance on a financial decision-making task; i.e., the Columbia Card Task (Figner, Mackinlay, Wilkening, & Weber, 2009). Students could earn a significant bonus to their base payment for completing the study. Further, we added an experimental manipulation of psychological stress (i.e., time pressure) during the decision-making task because previous research suggests that mindfulness is particularly beneficial in managing stress. The changes in the study design are expected to both increase sensitivity to the mindfulness intervention and improve the external validity of the study.

4.1.1.1 Sample and methods

Business Administration students (n = 80) were recruited at Erasmus University Rotterdam. All participants received one of the following equally long audio fragments: meditation with breathing (MF experimental groups) or listening to a chapter from Lord of the Rings (neutral condition). These audio clips were used successfully in previous studies (e.g., MF1). Following, they completed a series of questionnaires on mindfulness (Kentucky Inventory of Mindfulness Skills; KIMS), body vigilance (Body Vigilance Scale; BVS), and a financial risk-taking task (Columbia Card Task-Stress Version with a timer). In the Columbia Card Task, cards are presented face-down on the screen. Participants had to indicate how many cards they would want to turn over given specific loss and gain information. Turning over more cards both increased the gain but also increased the risk of losing all
the money gained in that round. If a single loss card is turned over, the participant would lose that round. Thus the subjects were continuously trading-off expected value vs. risk.

Students were divided into four groups using a standard 2x2 design for experimental research (No Stress & No MF; Stress & No MF; No Stress & MF; Stress & MF). The sample contained an equal number of men (n = 40) and women (n = 40) with a mean age of 21 years (SD = 2.38). They were offered a base pay of 5 euros, and could earn a bonus on top of this (as high as 12.50 euros in total).

4.1.1.2 Findings

Using ANOVA, we compared the participants who received a mindfulness exercise to the participants in the control condition. No differences were observed between the two groups for trait mindfulness (F(1,64) = 1.25; p = .27), which indicates that at the start of the study participants between the conditions did not differ in trait mindfulness. Compared to the control conditions (Stress/No Stress & No MF), however, ANOVA-analyses indicated that participants who received the mindfulness audio clip (Stress/No Stress combined for the purpose of this analysis) indicated a significant increase in body vigilance (F(1, 64) = 46.32; p < .01) and state mindfulness on both subscales of the TMS: decentering (F(1,79) = 6.93; p = .01) and curiosity (F(1, 79) = 7.79; p = .01). Thus, as expected participants in the mindfulness condition demonstrated higher levels of state mindfulness, but not in trait mindfulness.

A series of one-way ANOVAs showed that mindfulness (both Stress/No stress groups combined for the purpose of this analyses) showed that compared to controls, people who practised mindfulness took loss information significantly stronger into account when making a decision (F(1,63) = 8.63; p < .01), and decided to turn over less cards when presented with higher loss-amounts (F(1,63) = 9.29; p < .01). Thus, mindfulness helped participants pay more attention to loss information.

There was a significant trait neuroticism x mindfulness training interaction that explained 8% of the variance (p < .05). There were no other significant effects (e.g., non-significant effects of stress and stress x mindfulness). As shown in Figure 4-1, mindfulness training improved decision-making over time (smaller deviations from optimal solution) only for individuals high in trait neuroticism. 7Only in individuals high in trait neuroticism were the differences between groups (mindfulness, control) statistically significant for the deviation from the optimal solution, indicating that mindfulness training only helped people with high levels of neuroticism.

---

7 It should be noted that deviation from the optimal solution in the Columbia Card Task is represented by change scores in the deviation from optimal solution during the first block of trials in the game, compared to the last block of rounds in the game. The scores are an index of learning over time. The fact that all of the groups have negative scores means that everyone - on average - improved over time and demonstrated an approach to the optimal solution. However, only the high neuroticism individuals were affected more strongly by the mindfulness exercise and approached the optimal solution more strongly.
4.1.1.3 Implications for xDelia learning pathway design

Mindfulness improved optimal decision-making in participants with high levels of neuroticism, that is, people who are more subject to strong and changing emotions. So, people who experience high levels of stress may benefit most from mindfulness training. This is not surprising given that mindfulness is well-established in the field of health psychology, where it has known effects on dealing with stress. It is therefore reasonable to assume that people scoring high in neuroticism are characterized by high levels of stress and more likely to benefit the strongest from mindfulness training. Those with a high level of emotionality are more likely to experience stress whilst trading or be susceptible to the stressful impact of trading. Based on the present findings, these individuals would benefit most from undertaking mindfulness training.

The findings imply that mindfulness is particularly effective in people with high neuroticism. In people who score lower on neuroticism, and who may be less prone to stress or negative emotions, mindfulness skills may impact their emotion regulation skills less (e.g., because they already possess good skills on emotion regulation). It should also be noted, however, that in the current study we only used a brief exercise on mindfulness. According to the literature, mindfulness involves a set of skills which require intensive training (e.g. Segal, Williams, & Teasdale, 2002). Hence, the findings could also simply indicate that longer training periods are required for low neuroticism groups to experience positive effects on the relationship between emotion regulation and decision-making.

This study provided an important empirical foundation for testing the benefits of short mindfulness-based interventions on financial decision-making in the target populations (traders and investors) and confirmed that a more extended approach may be required to obtain more robust effects (e.g., repetition of a single exercise to investigate learning effects over time, or a training of several weeks). Results fed into the development of the xDelia Mindful Trading Training and study SM-7 and provided insight on the usefulness of mindfulness in target population (i.e., private investors), who may well be characterized by high emotionality during trading.
4.1.2 TIG-1: Initial evaluation of Two Index Game (EUR)

This study was run in addition to the series of studies previously outlined in D9-2.3.2.

Within xDelia, the Two Index Game (TIG) was developed as an index of the disposition effect and given high priority within Saxo Bank. The disposition effect is a common decision-making bias in traders and private investors, and is defined as the tendency to hold on to losing assets too long and cut winning assets short (Odean, 1998). The TIG is a simplified trading simulation game which mimics some elements of a trading task. Within xDelia, two versions were developed: TIG-diagnostic (which indexes the size of the disposition effect) and TIG-didactic (which in addition to indexing the disposition effect also provides a short training on how to avoid the disposition effect). This study focused on using the TIG-diagnostic. As we assume it may have important diagnostic characteristics with respect to the target population of xDelia, private investors. However, before implementing this newly designed tool in a series of studies, it seemed recommendable to test which players understood the game and could play it easily in a sample of novice decision-makers with respect to the trading context (i.e., students). Furthermore, this study tested whether the TIG is a valid tool to observe the disposition effect.

4.1.2.1 Sample and methods

Business Administration students (N =100) were recruited at Erasmus University, Rotterdam. All participants played the TIG Diagnostic version twice (Saxo Bank; v 0.7) to diagnose participants’ measure of disposition effect (DE). They played the games on the same lab appointment, immediately after each other. In addition, they completed a short survey afterwards on state emotion regulation (ERQ), experienced emotions (Body Vigilance Scale) and state mindfulness (Kentucky Mindfulness Inventory). For their participation, they received a financial remuneration which depended on gameplay, and was related to their ‘Distance from optimum’-score.

4.1.2.2 Findings

Overall, participants found the game engaging (66%). Participants indicated to understand the game well but also reported that it was difficult at times. In terms of gameplay, students understood the game well and were able to play it without problems.

The disposition effect scores were normalised by applying a log transformation. Scores under 0 represent a relative absence of the disposition effect in participants trading data, while all scores above 0 indicate the existence of the disposition effect. There was a large variability in the disposition effect scores. In our sample, we noted that 73.5 % demonstrated a score above 0. This finding suggests that the sensitivity at which the TIG detects the DE is appropriate given that the DE is proposed as a common bias in financial decision-making.

All players played the same game twice. The DE in the first game correlated well (test-retest $r = .55; p < .01$) with the disposition effect score in the second game. After controlling for state emotion regulation and state mindfulness (TMS), partial correlations show that the disposition effect scores between the first and second game correlate even stronger ($r = .78; p < .01$). This shows that the disposition effect (as indexed by the TIG) not only represents a relatively stable phenomenon in the short term but also indicates that to some extent state emotions contribute to the disposition effect and thus, that trading is associated with emotion regulation.

To investigate this assumption more closely, we divided the sample in high (n=51) and low (n=31) reappraisers (RA), high (n=50) and low (n=32) suppressors (SUP), and high (n=23) and low (n=59) mindful (MF) participants based on their scores in the state-ERQ (high RA/SUP: score 4 or higher; low RA/SUP: <4) and the state mindfulness scale TMS (high Mindful : scores 3 or higher, low...
A series of ANOVAs with within-subjects variable disposition score (in the first game and in the second game) and between-subjects variable high or low use of a strategy (RA, SUP, or MF) showed that only high suppressors demonstrated a significant difference compared to low suppressors on the difference in disposition effect between the two games ($F(1,79)= 5.08; p = .03$). Low suppressors demonstrated a significantly greater disposition effect in the first game compared to the second game. This indicates that emotion regulation strategies can financial decision-making and especially during periods of prolonged and potentially stressful trading. Nevertheless, the finding that only suppression proved beneficial to the disposition effect is surprising. In the literature (Gross & John, 2003), reappraisal strategies are commonly associated with more positive effects (such as on subjective well-being or stress), while suppression strategies only temporarily block the problem from the individual’s awareness (which may help dealing with daily life in the short term, but is associated with negative effects in the long term). Yet, it could be argued that suppression strategies are therefore more effective in acute and short periods of stress. The literature does show that suppression strategies generally exhaust individual cognitive resources quickly (e.g. Gross & Thompson, 2007) so it is not unlikely that when trading for a longer period (e.g., a full day), at some point suppression strategies will become counterproductive and reappraisal strategies appear more effective for long-term periods of trading.

4.1.2.3 Implications for xDelia learning pathway design

The study showed that trading in a trading simulation game (TIG) demonstrated biases in a group of players at a similar rate to that seen in trading activities of actual financial decision-makers.

The disposition effect was detected in a majority of our participants, and proved relatively stable over two play sessions that were completed consecutively, with no time lag in between each session. Hence, the DE can be detected in a simulated trading environment (using the TIG) at a rate equal to that found it in real world trading. An interesting implication of this finding is that a gaming environment can be used to investigate trading-specific biases like the disposition effect. It should also be noted that although this study used students, other studies within xDelia that used private investors yielded similar results (e.g., SM-8).

The present findings further suggest that the TIG could be a valuable tool for detecting an individual’s tendency to demonstrate the disposition effect, thus, providing private investors and traders with a tool that could demonstrate how prone they are toward the disposition effect. In addition to its diagnostic value, a didactic version of TIG has also been produced within xDelia. In this version, a short explanation on the disposition effect is given, after which the player can practice avoiding the disposition effect actively. It would be interesting to examine how learning about the disposition effect and dealing with it in that didactic version, could help actual trading behaviour. Yet, the findings from the current study show that TIG is a valid instrument to diagnose the disposition effect.

4.1.3 S-M1 and S-M2: Investor feedback trial (OU, Saxo)

This study evaluates the first phase of our learning pathway design (Diagnosis and Feedback) using a sample of study participants drawn from our primary target audience (investors trading their own portfolio on a regular basis).

Participants complete an online questionnaire on their habitual emotion regulation approaches. The online system also calculates a disposition effect score from their real trading history records. They

---

8 The disposition effect is the tendency to hold onto losing assets longer than winning assets.
receive structured personalised feedback on their emotion regulation approaches, their disposition effect and the likely meaning of these for their trading behaviour and performance. The automated online feedback (which follows completion of the questionnaire) is via a combination of text and video.

Following this feedback participants were offered an opportunity to get further feedback by playing an online game – the Two Index Game (TIG) diagnostic. The game provides feedback on disposition effect from the gameplay (a simple trading task).

Our evaluation encompasses both user experience and perceptions of the learning value and an examination of the behavioural effects of the intervention.

Research questions:
1. Do participants see positive learning value in the experience and do they find the experience engaging?
2. Does the TIG have diagnostic value? That is, does a disposition effect score in the game have predictive value in relation to propensity to display disposition effect in a real trading context?
3. Does the learning intervention influence behaviour? Are there measurable differences in participants’ behaviour relative to non-participants before and after participation in the learning intervention?

4.1.3.1 Sample and methods

A sample of 270 Saxo Bank clients were recruited to take part in the study and invited to continue on to play the TIG online; of these 270, 64 went on to play the TIG online (222 of the 270 agreed to share their questionnaire and disposition effect data for analysis by the research team). These participants (unlike our student samples) all have experience of real world trading using an online platform and placing their own money at risk. Participants varied in trading experience from 3 months to 32 years and in age from 22 to 67.

All participants completed the emotion regulation survey and received structured feedback on their emotion regulation approaches and the extent to which their trading record showed them to have a disposition effect.

A follow up survey was administered to participants to gain feedback on their perceptions of the learning value of the diagnosis and feedback process.

Disposition effect scores were also calculated for all participants in a follow-up period of 11 weeks following the intervention.

4.1.3.2 Findings

4.1.3.2.1 Two Index Game as a diagnostic instrument

A key finding is that a disposition effect score in the TIG shows a significant association with a disposition effect score in real world trading. Disposition effect scores in actual trading and in the TIG show a significant correlation (0.28, p<.05). When we use the TIG DE score to predict whether or not participants show a disposition effect in their real world trading (based on binary logistic regression), the TIG score correctly classifies 67% of participants – significantly better than chance.

9 See D18-2.4.2
This correlation between TIG DE score and trading DE score related to actual trading remains consistent when trading DE is measured at three different time periods.

This finding is especially notable since we know from prior work that disposition effect can be highly influenced by context. Participants were, in their real world trading, trading a wide range of different assets and facing dissimilar and changing market conditions. That we find a relationship between a game based measure (where the only incentive is a desire to get a high score) and a real world trading context with significant sums of money at stake is highly notable. Particular implications include:-

- Strong support for the view that the disposition effect is underpinned by individual characteristics as well as trading context.
- Support for the value of the TIG as a diagnostic tool and training environment for addressing the disposition effect.

We also found a relationship in the TIG gameplay between disposition effect and lower game profits, consistent with our expectation that when trading on the basis of some level of information a disposition effect reduces average trading performance.

4.1.3.2.2 Participant perceptions of the learning value of the intervention

The majority of respondents were positive about the learning value of the intervention.

*Emotion regulation diagnostic questionnaire:* Most felt that the emotion regulation feedback and materials were informative (76% agree, 10% disagree); caused them to learn something new (68% agree, 12% disagree); were interesting and engaging (80% agree, 10% disagree) and provided a good basis for reflective learning (78% agree, 6% disagree).

*Disposition effect feedback and learning materials:* Most felt the feedback and learning materials on the disposition effect to be informative (82% agree, 9% disagree); caused them to learn something new (76% agree, 12% disagree); were interesting and engaging (82% agree, 9% disagree); gave them insight into their own trading (85% agree, 3% disagree); and provided a good basis for reflective learning (76% agree, 6% disagree).

*Two Index Game:* The group who went on to complete this game were also mostly positive about the experience. They felt it was fun (69% agree, 12% disagree); very few felt it was boring (4%) and likely, in combination with other learning elements to help them learn to reduce their disposition effect when trading in the real world (58% agree, 8% disagree).

4.1.3.2.3 Relationship of questionnaire scales to trading behaviour

We hypothesised a relationship between trait emotion regulation (as measured by the ERQ) and disposition effect. What we found was more complex. Neither reappraisal nor suppression showed a direct correlation with disposition effect (the tendency to realise gains more readily than losses). However a habitual reappraisal approach to emotion regulation was associated with greater readiness to realise both losses and gains, as opposed to a difference in the balance between realising gains and losses. This is contrary to a parallel study carried out at OUBS (Richards, unpublished PhD thesis) which did find reappraisal associated with lower disposition effect and to the finding in another xDelia investor study (S-M8) where we found reappraisal to be associated with lower disposition effect in the TIG and suppression associated with higher disposition effect.

We consider there to be two plausible explanations of this result. A) Traders may be more likely to answer positively to questions about emotion regulation if they experience significant need to regulate emotions during their trading. In other words the ERQ as used here may be picking up both approach to emotion regulation and amount of regulation, thus confounding quality of regulation.
with need to regulate. B) Reappraisal may be associated with different goals. If the goal is to manage emotional reactions in the interest of a balanced approach to decision-making then we would expect a reduction in disposition effect. However if the goal is to avoid negative emotions by avoiding responsibility for outcomes then it could lead to greater disposition effect. It is also possible that this result reflects a lack of relationship between emotion regulation and disposition effect. What is more likely, given the different findings in other studies is that there are contingencies we are not accounting for.

These results suggest that the online diagnostic questionnaire should be developed further to improve validity in relation to trader behaviours. First, we suggest recasting the ERQ questions to focus on habitual approaches to emotion regulation in everyday life. Second, we suggest developing a dedicated scale which measures traders own perceptions of susceptibility to emotion effects on their trading. (This work is now ongoing).

### 4.1.3.2.4 Effects of intervention on trading

We calculated disposition effect scores\(^{10}\) for all Saxo clients with sufficient trading history in 3 world-regions prior to recruiting for our feedback trial (Time 0). All Saxo clients in those regions who had sufficient trading history to calculate a reliable disposition effect score were invited to participate in the learning trial via Saxo Bank’s standard marketing email channel. The initial disposition effect calculation was carried out on between 3 and 18 months of trading history (depending on client longevity) at time 0. We then calculated DE scores from a time 6 weeks later over 14 weeks (time period 1) and from 11 weeks after t0 over 7 weeks (time period 2). The training intervention occurred between time 0 and time 1. Due to the time taken to complete surveys and the TIG the first period contains some days trading for some participants prior to participation (although most after). The second time period has no overlap but includes data from a smaller sample and has DE calculated over a shorter period. Thus we examine both changes from t0 to t1 and from t0 to t2.

<table>
<thead>
<tr>
<th>Number still trading in each time period</th>
<th>did survey</th>
<th>no survey</th>
<th>Sig. difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>t0</td>
<td>222</td>
<td>5269</td>
<td>100.00%</td>
</tr>
<tr>
<td>t1</td>
<td>100</td>
<td>1561</td>
<td>37.04%</td>
</tr>
<tr>
<td>t2</td>
<td>68</td>
<td>939</td>
<td>25.19%</td>
</tr>
</tbody>
</table>

Table 4.1 – Drop out over time

Drop-out rates are both a problem for analysis and a source of data. We deal with the analysis problem by using a multiple imputation approach for missing values which adjusts significance of findings for error of estimation (see below for an explanation of this approach).

The dropout rate also provides an interesting finding (Table 4.1). If we take missing data at the time of second and third DE measurements as a proxy for ceasing trading, then those that took part in our study are significantly more likely to be still trading with Saxo at the end of a five month period.

\(^{10}\) Disposition effect scores are calculated by first calculating the proportion of all trading gains which are realised (i.e. number of gains realised over total holding time for gains) and the proportion of trading losses which are realised in the period. The disposition effect score is calculated as the proportion of gains realised over the proportion of losses realised. The natural log of the ratio is taken to normalise the distribution. This also results in scores above zero representing a disposition effect and scores below zero representing an inverse disposition effect.
This could be a self-selection effect, but nonetheless it provides some support for the value of this kind of engagement with bank clients.

When we examine change in DE over time, we find that DE for participants and non-participants is not significantly different between cohorts at time 0 but shows a significant (p<0.05) difference at time periods 1 and 2. Figure 4-2 shows estimated means over time by cohort.

![Figure 4-2: Study S-M1/2: Mean disposition effect by time and cohort](image)

<table>
<thead>
<tr>
<th></th>
<th>Time 0</th>
<th>Time 1</th>
<th>Time 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not a participant</td>
<td>0.52</td>
<td>0.49</td>
<td>0.58</td>
</tr>
<tr>
<td>Participant</td>
<td>0.43</td>
<td>0.21</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Our ability to separate out effects of feedback from survey and actual trading from engaging with the TIG online is limited by the smaller number of TIG participants.

The chart below (Figure 4-3) separates out those who participated in the TIG. The differences (given smaller sample size are not significant but as can be seen this does suggest that further decrease in DE between period 1 and 2 may possibly be associated with engagement with the TIG).
Appropriate strategies for handling missing values in data depend on the nature of the “missingness”. If missing values are randomly distributed, with no relationship to the value of the missing variable or to other variables in the data set (known as missing completely at random or MCAR), then both listwise and pairwise deletion of missing values will yield unbiased estimates when analysing the data, although, in the case of pairwise deletion, standard errors will be biased towards zero and listwise deletion will be inefficient, because it discards observations. Multiple imputations likewise yields unbiased estimates of parameters, but is more efficient than listwise deletion, and gives unbiased estimates of standard errors. A less stringent assumption concerning missing data is that they are missing at random (MAR). That is, if \(x\) is a variable with missing (i.e., unobserved) values and \(x_{\text{mis}}\) is a variable coded 1 when \(x\) is missing and 0 otherwise, then, after controlling for other observed variables in the analysis, \(x_{\text{mis}}\) is independent of the value of \(x\). Where data are MAR but not MCAR then both pairwise and listwise deletion will yield biased estimates for parameters. In contrast, multiple imputation yields unbiased, fully efficient estimates for parameters and unbiased estimates of standard errors. Where \(x_{\text{mis}}\) is related to the underlying value of \(x\), after controlling for observed values, this is known as non-ignorable (NI) missing data. Where data are NI missing, pairwise deletion, listwise deletion and multiple imputation all yield biased estimates of parameters (Allison, 2001; King, Honaker, Joseph, & Scheve, 2001).

Although it is possible to distinguish empirically between MCAR and MAR using observed data, it is not generally possible to determine from the observed data whether data are NI missing. Since there is no data to determine whether \(x\) and \(x_{\text{mis}}\) are correlated. However, in this case while we cannot examine the correlation between \(x\) at time 2 (3) and missingness at time 2 (3), a good proxy is to examine \(x\) at time 1 and missingness at time 2 and 3. The observed correlations are both very small and non-significant offering reassurance that \(x\) is MAR not NI. We also find that \(x\) at time 0 is not significantly correlated with participation in the study.

Note on missing value treatment

*Figure 4-3: Study S-M1/2: Mean disposition effect by cohort and time (separating TIG cohort)*

<table>
<thead>
<tr>
<th></th>
<th>Time 0</th>
<th>Time 1</th>
<th>Time 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non participants</td>
<td>.5175</td>
<td>.4871</td>
<td>.5724</td>
</tr>
<tr>
<td>Survey but not TIG</td>
<td>.4551</td>
<td>.2516</td>
<td>.2644</td>
</tr>
<tr>
<td>Survey and TIG</td>
<td>.1733</td>
<td>.1616</td>
<td>-.0533</td>
</tr>
</tbody>
</table>
Thus the most appropriate approach to dealing with the missing data is to use multiple imputations. We used the approach described by King and associates (King et al., 2001). The missing values are imputed multiple times (usually 3 to 5 is sufficient) to generate multiple data sets. Analysis is then carried out on each imputed data set separately. The missing value imputation was carried out using SPSS, with 5 imputed data sets. Model parameters are calculated as the average of the parameters from each separate analysis. This process gives unbiased estimators for model parameters and allows an adjustment to standard errors to allow for the uncertainty of imputed values. Standard errors are calculated by:

\[
SE(q)^2 = \frac{1}{m} \sum_{j=1}^{m} SE(q_j)^2 + S_q^2 (1 + 1/m),
\]

where \( m \) = number of imputed data sets, \( SE(q_j) \) is the standard error of the \( j \)th estimate of parameter \( q \),

\[
S_q^2 = \sum_{j=1}^{m} (q_j - \bar{q})^2 / (m-1).
\]

### 4.1.3.3 Implications for xDelia learning pathway design

This study suggests that our diagnosis and feedback approach is well received by the target group and provides a good basis for reflective learning. The TIG is not only well received by learners in the target group, but also shows promise as a valid diagnostic instrument and learning environment, given the significant relationship between behaviour in the game and behaviour in real world trading.

Perhaps most importantly there is evidence of an effect of this intervention on actual trading behaviour as measured by DE before and after the learning intervention.

The study also suggests a need for further development of the online diagnostic questionnaire.

### 4.1.4 S-M3: Emotion regulation training by means of an arousal-based game (EUR)

Following a previous study in a small sample of investors who tested the Aiming game (IS 3), development of the game continued. The next version of the game deviated from the original version in so far that it used a different psychophysiological input signal (ECG) from the previously used EPOC EEG signals as an established measure of arousal (heart rate) to influence gameplay performance, it was situated in a more neutral game context (a space environment) and it was renamed as the Space Investor game (SI).

The major research question for this experiment was whether emotion regulation could be trained using this game and to determine which future modifications would be needed. Therefore, we relied heavily on the user feedback provided by the participants to feed into development of a new version of the game.

### 4.1.4.1 Sample and methods

All participants (n=32) were (International) Business Administration students. Most of the participants were male (n = 24; 75 %) with a mean age of 20.35 (SD = 2.60). Participants played the game, completed a survey and received a small financial reward for participation in the study. Participants were recruited via an online bulletin board where they could sign-up for an available timeslot. Upon arriving at the lab the experimenter provided a brief introduction to the experiment. Then, the experimenter attached the heart rate sensor to the participant (a small wireless box attached D9-2.3.3 – In-depth studies: results (Year 3)
to a strap on Velcro belt) before they played SI for approx. 25-30 minutes. The arousal bar was inactive for participants in the control group (N = 14) and active for the participants in the treatment group (N = 17). Finally, participants received a small survey asking about their gameplay experience (i.e., how difficult they found the game to play, whether they understood the game, what they thought about the graphics, whether they had any suggestions for improvements). One person did not complete the questionnaires due to time constraints and was excluded from our analyses.

4.1.4.2 Findings

The main findings from this study were: a) SI induces arousal in several game levels, b) arousal management helps players to perform better in SI, and c) SI requires further tweaks to strengthen the game, such as using a continuous rather than a discrete arousal input.

Overall, 71% of all participants indicated that they liked the idea behind the game and considered it to be a viable approach toward training emotion regulation. SI was built so that each level would be more difficult than the last. We expected that each level would thus induce stronger levels of arousal and give the participants greater opportunity to practice emotion regulation strategies. Figure 4-4 shows that all of the levels where the arousal bar is active (levels 2 to 5) demonstrate higher levels of arousal than the baseline level (level 1). Furthermore, it can be seen that level 4 could be tweaked to ensure that each level is perceived as being increasingly more difficult and arousing than the last level played.

The findings point in the expected direction but there were no statistically significant differences in self-reported arousal levels between people who played it with biofeedback active and who played SI without biofeedback ($F(1,29) = .37; \ p = .12$). However, the sample size was extremely small, so these results should be interpreted with caution.

![Figure 4-4: S-M3: Self-reported arousal (SAM) at increasing game difficulty levels.](image)
The findings show that the ability to manage arousal varied between levels and conditions. Most participants (77%) reported that SI actually was easy to play and that they experienced low levels of emotions/arousal as a result of playing the game. Accordingly, a repeated measures Analysis of Variance (MANOVA) demonstrated that there were no differences between experimental and control participants on performance - the number of asteroids destroyed per level ($F(4,26) = .90; p = .48$). Yet, as Figure 4-5 shows, there does appear to be a peak difference in number of asteroids destroyed, but only for the level with the highest observed arousal levels (i.e., level 5). These findings can be taken to indicate that levels 2 to 4 require evaluation in terms of difficulty to ensure that these generate enough arousal and to bridge the gap in difficulty between the first levels compared to the final level (5) which is perceived as more difficult.

![Figure 4-5: S-M3: Number of asteroids destroyed between experimental and control group](image)

Further attesting to this view, we ran a series of correlations between game statistics and self-perceived ability to regulate emotions during the game for participants in the treatment group. We observed that only level 5 showed significant correlations between self-perceived arousal management capability (‘did paying attention to arousal help you perform during the game?’) and the number of asteroids destroyed ($r = .51; p < .01$) and the number of resources collected ($r = .51; p < .01$). Yet, no differences were observed in the actual ability to shoot the asteroids between groups in that level ($t(29) = 1.08; p = .29$). Although this likely reflects a lack of power due to small samples sizes, this remains open for future research.

In support of the idea that emotion regulation can be trained through biofeedback games like the SI, 65% of people who played it with biofeedback active reported that paying attention to their arousal levels helped them to play the game better. Indeed, in the experimental condition, paying attention to arousal was associated with self-reported use of the arousal bar ($r = .59; p < .01$).

Following these comments, a more continuous arousal bar computation was provided by FZI and difficulty levels along with several other minor aspects of the game were adjusted.
4.1.4.3 Implications for xDelia learning pathway design

The study has shown that SI is an effective space where participants in the learning intervention have the ability to monitor their progress or practice their newly learned skills on emotion regulation. Findings imply that SI can help people to train emotion regulation strategies within a game environment and several future iterations have been recommended, such as combining the biofeedback signals with explicit tips and hints on strategies that participants can use during the levels.

Unfortunately, the installation of the required sensors for the game took longer than anticipated due to a number of technical difficulties (e.g., involving unreliable Bluetooth communication between sensors and computers). This caused a marked delay in the implementation of the study. Further, these delays impacted the design and implementation of the next batch of studies, in particular SM-7 as we planned another iteration round following calibration study S-M3 and the longitudinal study using SI as a training approach (S-M7). Yet, by prioritising the major elements to be tweaked in SI and communicating frequently with various partners (e.g., game developers from BTH), a new version was delivered in time for S-M7 to proceed with a limited delay. Moreover, a number of other technical issues were resolved in the process.

However, as we only addressed the major issues in SI during the iteration cycle for S-M7, several features and tweaks could benefit the approach. Some of these have already been addressed for SM7, or later on for studies involving partners within the project (e.g., studies by OU in S-M8). Other issues, however, remain open for future research. To help people train emotion regulation strategies it could help to provide hints and tips during gameplay, which they can then practice in the game environment. Interestingly, people indicated that they used two strategies automatically for emotion regulation: 1. Controlling breathing, 2. Focusing and staying calm. The game, however, does not make any explicit claims or suggestions on how to regulate emotions. Implementing tips and hints into SI could help facilitate its learning effects. Another implementation would be to remove some of the redundant variables (e.g., explanations on lasers you can’t buy) and explore the effectiveness of the difficulty adjustments following S-M3.

To conclude based on the present findings, several changes could improve the SI game further. Nevertheless, the current version of SI is a promising training tool to help people manage their (emotional) arousal levels.

4.1.5 S-M4: Improvement of Emotion Regulation by playing the Auction Game V2 (FZI-IPE)

Research questions: Three questions were considered: (1) Does the Auction Game – Didactic (AUC-DID) provide an effective space for practicing emotion regulation? (2) Are different ER strategies (measured by physiology and ER questionnaires) reflected in different game performance? (3) Do players improve their skills in emotion regulation (measured by physiology and ER questionnaires) in different modes in the game?

The first part of the study aimed to calibrate the AUC-DID and the influence of biofeedback on emotion regulation / decision performance. The second part of the study was conducted to examine

---

11 Difficulties with Bluetooth communication between computers and sensors have been a consistent feature of these studies. Problems only really became apparent as we started running multiple installations of the games in parallel. Problems become more likely the greater the number of installations run in parallel. Problems were often also highly variable between settings suggesting that they may be also related to background Bluetooth activity.
the training effect of biofeedback and arousal-influence towards emotion regulation and decision performance. The effect of emotion regulation training intervention on other economic tasks was covered by studies S-M5 and S-M6.

The AUC-DID was designed for providing decision-makers with a training tool, which penalises the inability to regulate one’s own level of arousal. More specifically, in the current version of the game, the setup is designed to penalise high levels of arousal and the inability to down-regulate arousal. In order to do so, the game has to meet two major requirements. First, the game must be capable to actually induce high levels of arousal. Second, the game must significantly penalise subjects that have high levels of arousal and those subjects that are unable to practice emotion regulation. On the other hand, it must reward those that are able to down-regulate high levels of emotional arousal. For that reason the variance of the price signals around the true mean, and hence the difficulty of estimating the mean, is connected to the individuals’ level of arousal. Therefore, the physiological level of arousal was positively correlated with the game difficulty.

In addition to using the emotion regulation categorization of Gross and John (2003), we also directly asked the subjects of study part I which strategies they applied in order to down-regulate their emotional arousal and analysed the effectiveness of these strategies. It was found in previous research that better emotion regulation goes along with higher heart rate variability (HRV), because a higher HRV can be interpreted as a more sophisticated interplay of the parasympathetic and sympathetic branches of the autonomous nervous system (Appelhans & Luecken, 2006; Sütterlin, Herbert, Schmitt, Kübler, & Vögele, 2011). In addition to evaluating the game, the two studies were also designed to provide further insight into emotion regulation processes and how they are manifested in the decision-making process.

![Biofeedback (BF) vs No Biofeedback (NBF)](image)

Figure 4-6 shows a screenshot of the two treatment groups of study I; with and without biofeedback (changing colours of clouds and arousal meter). This depicts a game screen of the ‘Biofeedback’ (BF) treatment (left) and the ‘No Biofeedback’ (NBF) treatment (right) in the for study part I. In the BF treatment the players get a continuous visual feedback on their own current emotional state in the form of an arousal bar at the top right corner of the screen and the changing colour of the clouds. This feature is aimed at drawing the player’s attention to the emotional state and interoception.

We also investigated whether the sensitivity for emotion regulation along with the display of biofeedback had an impact on game performance. In the second part of the study, again we employed a between-subjects design with two treatment groups. In the no influence (NI) treatment, the players played the game in a mode where neither the emotional state (via clouds changing colour or arousal meter) was indicated nor did it influence the difficulty of the gameplay. More precisely, the
arousal had no influence in this mode of the game. Also, for this group the instruction material differed slightly: Participants in the NI treatment only received the information that “paying attention is an important task to regulate your emotions,” without referring to arousal. In contrast, participants in the biofeedback treatment (BF) were given information about how game difficulty would increase with physiological arousal. Therefore, it was especially interesting to analyse, whether an increased sensitivity for emotion regulation in fact boosts decision performance in the AUC-ER.

4.1.5.1 Sample and methods

Both experiments were conducted at the laboratories at FZI. Study part I was conducted in late 2011 and study part II in early 2012. Both studies were based on between-subjects designs, i.e. subjects only experienced one of the treatment conditions. The ORSEE software environment was used to recruit participants from a pool of university students (Greiner, 2004). Table 4.2 provides a summary of the core parameters of the two experiments.

In study part I, 36 subjects participated. Subjects received detailed instructions on the game and how their physiological states interact with game difficulty. In addition to that, subjects in the BF treatment got information explaining that the arousal bar and the changing colour of the clouds would serve as an indicator of their emotional state. Moreover, subsequent to the experiment, a questionnaire was used to collect information about subjects’ perception of the game and which emotion regulation strategies they used during the game. In the BF treatment, physiological arousal had an influence on game performance and players were provided with a biofeedback (N=44). In the NI treatment the game was designed so the physiological state had no influence on game difficulty whatsoever (N=24).

<table>
<thead>
<tr>
<th></th>
<th>Number of Participants</th>
<th>Physiological Measurement</th>
<th>Biofeedback</th>
<th>Arousal influence on Game Difficulty</th>
<th>Age</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part I</td>
<td>36</td>
<td>BF treatment: 19</td>
<td>YES</td>
<td>YES</td>
<td>23.39</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NBF treatment: 17</td>
<td>YES</td>
<td>NO</td>
<td>[20-28]</td>
<td></td>
</tr>
<tr>
<td>Part II</td>
<td>68</td>
<td>BF treatment: 44</td>
<td>YES</td>
<td>YES</td>
<td>22.06</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NI treatment: 24</td>
<td>YES</td>
<td>NO</td>
<td>[18-27]</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2 – SM4 design overview, where NI treatment did not receive any arousal information, nor did arousal influence game difficulty. Therefore NI refers to the baseline treatment

In both study parts subjects were compensated according to their performance in the game. In the first part we had a constant conversion of game €s earned into real cash, which was paid directly after subjects had finished the game. In the second study part the game was played within a set of games and the payment was commensurate to performance in one of these games, which was chosen at random. However, in both studies subjects had a high incentive to perform to the best of their ability in order to earn money.

Subjects were randomly seated in the laboratory and the use of the sensors was explained at the registration. Subjects received an envelope which contained the number of their working station and their identification number. To reduce distractions participants were sat in a cubical and asked to wear headphones. After the instructions were read via headphones to the subjects they started playing a tutorial which walked them through the game mechanisms. After completing the tutorial successfully subjects were given a 5 minute period to rest before they were allowed to proceed with
the game. At the end of each session, subjects were asked to fill out the Emotion Regulation Questionnaire (Gross, 1998) and the Risk Aversion Questionnaire (Holt & Laury, 2002).

4.1.5.2 Findings

We first focus on the evaluation of the AUC-ER. The physiological data of our experiments showed that the game in fact induces high levels of arousal: the average arousal level for a subject, which can take on values between 1 and 5, was 3.272 with a standard deviation of 1.285. For almost all subjects (97%) the highest arousal level 5 was reached at some point of the game, and for most subjects (65%) all possible values from 1 to 5 were recorded during gameplay. After the first study part most subjects reported that they liked the game and also felt highly aroused.

Subjects were receptive toward the concept of regulating their emotions in order to ease the game difficulty. Since the game goal was to down-regulate high levels of emotional arousal in order to perform well in the game, it was important that we could actually connect decision performance to arousal. As depicted in the graph on the left in Figure 4-7, the subjects’ average decision performance decreased significantly with increasing levels of arousal. In the analysis we define the average decision performance as the percentage of how many decisions were taken correctly, because this is the most fine-grained performance indicator here. This is also supported by a Pearson’s correlation coefficient of (N=104, Pearson’s \( r = -0.359, p < 0.001 \)).

![Figure 4-7: S-M4: Subjects average decision performance depending on their arousal state.](image)

Hence, the two main requirements for an emotion regulation learning environment are met: the game induces high levels of physiological arousal and the arousal in turn significantly worsens an individual’s decision performance.

In summary we can conclude that the AUC-DID rewards active down-regulation of high levels of arousal and provides a learning environment in which emotion regulation can be practiced and rewarded. Moreover, the results show that the players’ average arousal levels during the game are correlated with the average money earned in the game (N=104, Pearson’s \( r = -0.377, p < 0.001 \)), the level which was reached (N=104, Pearson’s \( r = -0.363, p < 0.001 \)), and the average number of decisions taken correctly (N=104, Pearson’s \( r = -0.331, p < 0.01 \)). Correspondingly, the graph on the right in Figure 4-7 shows that the average arousal level is lower for those subjects who manage to successfully finish the game up to the last level (N=20) compared to the rest of the subjects.

Subjects that apply either the emotion regulation strategies ‘emotion reappraisal’ or ‘emotion suppression’ performed better in the game than those subjects who did not apply any of these strategies. In order to assess whether the subjects performed ‘emotional reappraisal’ or ‘emotional suppression’, the previously described ERQ was used (Gross, 1998).
4.1.5.2.1 Relationship of Emotion Regulation Questionnaire and behaviour

Based on the ERQ, we compute for each subject one score for emotional reappraisal and one score for emotional suppression. Both scores can take on values between 1 and 7. A subject is classified as a reappraiser, if his or her emotional reappraisal score is greater than 4. Equivalently, a subject is classified as a suppressor, if his or her emotional suppression score is greater than 4. With this classification, 36 subjects were identified as reappraisers and 11 as suppressors—the other 57 subjects are allocated to the group which does not predominantly use one of the two emotion regulation strategies. The results from the experiments show, that those subjects, who were classified as suppressors or reappraiser using the ERQ, achieved a better decision-performance than the rest of the subjects not classified as suppressors or reappraisers. Applying emotional reappraisal, which is cognitively demanding, results on average in a better decision performance in the AUC-DID than not applying emotional reappraisal (.712 vs. .661, N=104, t(102)=2.102, p<.05). Interestingly, also applying suppression strategies results in more beneficial financial decisions in the game compared to not using suppression strategies (.758 vs. .670, N=104, t(102)=2.354; p<.05).

Hence, both emotion regulation strategies seem beneficial for decision performance in the AUC-ER. Moreover, it is interesting to note that those subjects who neither apply emotional reappraisal nor emotional suppression fail more often to take any decision and hence get a 5.00 penalty. The results show that reappraisers and suppressors perform better in the game. Hence, the game rewards the use of these emotion regulation strategies.

This result provides further evidence for the assumption that effective emotion regulation may not only be beneficial for mental and physical health, but also for financial decision-making.

4.1.5.2.2 Relationship of Emotion Regulation measured by HRV and behaviour

Finally, we investigate whether differences in heart rate variability (HRV) are reflected in the bidders’ levels of arousal. As mentioned above, heart rate variability is a physiological parameter that can be interpreted as a proxy for emotion regulation capabilities (Appelhans & Luecken, 2006; Süttelins et al., 2011). The HRV LF/HF-ratio reflects the balance between the sympathetic and parasympathetic autonomous nervous system. We find a positive correlation between subjects’ LF/HF-ratio and their variance of arousal values during the game (N=104, Pearson’s r=.230, p<.05). We interpret this finding as an indication that subjects with a low LF/HF-ratio are more flexible to constantly keep their arousal down, and hence expose a lower variance on the arousal signal. However, we did not find a significant correlation between decision performance and heart rate variability.

In study part I we investigated whether an increased awareness of one’s own emotional state increased decision performance in the AUC-ER. Therefore, in the biofeedback (BF) treatment, the current level of physiological arousal is visualized to the players by means of an arousal meter and the changing colour of the clouds (left part of Figure 4-7), which ranged from green (low arousal), yellow (medium arousal), to red (high arousal). In contrast, there is no visualization of arousal in the no biofeedback (NBF) treatment. All subjects received the same instructions on how emotional arousal influenced game difficulty, except that the NBF version did not get any information on the arousal meter and the changing of cloud colours, since these elements did not occur in this game version.

It is interesting to observe that subjects of the BF treatment did not think that the arousal indication helped them; either to become aware of their current emotional state, or to help them with coming to better decisions. However, the average arousal for subjects in the BF treatment was lower and a t-test revealed that subjects achieved a better decision performance than the subjects of the NBF treatment (.65 vs. .74, N=36, t(34)=2.530, p<.05).

4.1.5.2.3 Participant perceptions of the learning value of the intervention

The results from this study demonstrate how an information system enriched with biofeedback can boost subjects’ awareness of their emotional state and increase decision performance. Moreover, in Part I we used a free text questionnaire in order to ask the subjects what kind of strategies they used
to actively down-regulate their arousal. Out of the 36 participants, 21 stated that they tried to regulate their breathing in order to regulate their arousal. Common answers were: “I tried to breathe calmly” and “I concentrated on breathing.” Two subjects tried to close their eyes when they realized that they were too aroused, and six subjects tried to “think of something different.” These strategies are very similar to professional meditation techniques as, for instance, applied in mindfulness interventions (cf. Kabat-Zinn et al., 1992).

Those subjects who applied active-breathing had on average an arousal level (and therefore measured heart rate) that was 0.5 lower than the arousal level of the rest of the subjects. Moreover, active-breathing and closing-eyes lead to increased average decision performance. However, the thinking-of-something-else strategy was not successful, as it increased average arousal values and reduced decision quality Subjects that stated in the questionnaire that they actively turned away from trading when they were highly aroused reveal both significantly lower average arousal levels (N=36, Pearson’s r=-.477, p<.01) and significantly higher average decision performance (N=36, Pearson’s r=.533, p<.01). As a side result we find that subjects that experience high subjective anger, when coming to a bad decision, also on average exhibit higher arousal levels (N=36, Pearson’s r=.374, p<.05). We conclude that the subjects picked up the game concept and pursued high effort in order to regulate their arousal level. In fact, even though most of the subjects are initially unfamiliar with the concept of emotion regulation, they actively tried to manage to down-regulate high levels of arousal.

Study part II was designed for investigating how conscious emotion regulation influences the subjects regarding their overall decision performance in the AUC-ER. To this end, we conducted a between-subjects experiment with two treatments. In the biofeedback BF treatment, the subjects received direct biofeedback and their arousal level had an influence on game difficulty. In the no influence (NI) treatment, the biofeedback was neither displayed, nor had it an influence on game difficulty. The results of the experiment show that subjects in the BF treatment performed better in the AUC-DID than those subjects who were in the NI treatment (.629 vs. .688, N=68, t(66)=-1.985, p=.051). Thus, we can conclude that conscious emotion regulation results in an overall better decision performance in the game. Recall that the influence of biofeedback can only make the game more difficult. Therefore, subjects in the NI treatment played a game in which the trading task is never more difficult than it is in the BF treatment. Therefore, it is particularly interesting to observe that even though the game was potentially more difficult in the BF treatment, the subjects still managed to achieve on average a better decision performance. In other words, those subjects that knew that their arousal level influenced game difficulty performed better in the game even though playing a possibly more difficult game. Moreover, also the overall arousal level was significantly lower in the BF treatment than it was in the NI treatment (.629 vs. .688, N=68, t(66)=3.820, p<.001). In summary, we conclude that biofeedback increases interoception and that conscious emotion regulation can in fact help to reduce levels of high arousal.

4.1.5.3 Implications for xDelia learning pathway design

The AUC-DID – with slight modifications – is well suited for the longitudinal study to (1) make subjects aware of arising emotions and (2) to effectively train to down-regulate high levels of arousal as a route to improved interoception and thus improved emotion regulation.

4.1.6 S-M5: Auction game as feedback tool on phasic physiological responses (FZI-IPE)

The goal of this study is to evaluate the Auction Game (AUC) as an assessment tool in relation to emotional overweighting of losses.

Research Questions: (1) Is effective emotion regulation manifested in higher decision performance and different physiological reactions? (2) Are physiological reactions in the game related to disposition effect measured with the TIG? (3) Could these measures be used as individual feedback?
One part of the study (part I) was elaborated in order to calibrate the AUC-DIAG and the Two Index Game (TIG). In more detail, the first goal of study S-M5 was to evaluate the AUC-DIAG as a tool to assess subjects’ emotional imbalance to the reaction of gains and losses on an individual basis. The emotions in this game (which were joy, frustration, regret and pride) were triggered via certain feedback information. These emotions could already be shown in a variety of studies to influence subjects’ decision process (c.f. Loewenstein & Lerner, 2003; Astor et al., 2011).

The second part of the study (part II) was conducted with the goal to examine the influence of a learning Intervention (AUC-DID and Space Investor (SI)) on phasic emotional responses in the AUC-DIAG and effects on the Disposition Effect (DE) as measured by the TIG. Moreover, the study used the TIG to investigate whether the propensity to show the disposition effect would also be manifested in physiological responses to joy and frustration in the AUC-DIAG.

4.1.6.1 Sample and methods

Study part I was conducted in late 2011 and study part II in early 2012. The ORSEE software environment was used to recruit participants from a pool of university students (Greiner, 2004).

Part I of the study consisted of 36 subjects. Subjects’ performance in a preliminary version of the TIG was measured. However, with respect to the AUC-DIAG, due to technical problems the laboratory set-up for measuring physiological responses to gains and losses only worked for 22 subjects. Part II of the study was carried out in combination with S-M6. HR was assessed for 81 subjects. The data was used to analyze how ER training affected subjects’ physiological responses in the AUC–DIAG. The physiological responses were then also employed as an indicator for subjects’ exposedness to the DE in the TIG. Finally, subjects were additionally asked to complete a questionnaire. Participants were compensated in cash according to how well they performed in both games. Participants were told in advance of taking part in the study that they would be paid this money to assure that subjects performed to the best of their ability.

4.1.6.2 Findings

4.1.6.2.1 Participant perceptions of the learning value of the intervention

Subjects liked both the AUC-DIAG and the TIG. They stated that they experienced high levels of joy and frustration during the games — both in free text fields but also in a subsequent questionnaire. They also stated that the games were highly engaging.

Figure 4-8 depicts the subjects’ average physiological responses to gains and losses. This is done by subjects’ successive drop in heart rate, directly after the experience of a gain or loss. As can be seen, the physiological response to winning and losing varies drastically. This provides more evidence that the game is highly arousing. This reflection of information (with positive or negative valence) in heart rate is consistent with other examination of psycho-physiology (cf. Bradley, Miccoli, Escrig, & Lang, 2008). However, given the small number of observations (N=22) and the high variance, inherent in psycho-physiological measures, we cannot analyse the data in a statistically valid way, with the goal to match the imbalance of physiological responses in the AUC-DIAG on the DE effect measured in the TIG.
4.1.6.2 Short learning intervention effects on ER and physiology in the AUC DIAG and the TIG

In Part II of the study we examined whether a short learning intervention of emotion regulation (with the AUC-DID (N=41) and/or SI (N=23) is manifested in the AUC-DIAG or the TIG. The other goal included trying to detect a correlation between the TIG and the AUC-DIAG and to potentially use the phasic responses in the game as a predictor for subjects’ individual physiological disposition.

Altogether 81 subjects took part in this second study part: 61 subjects were assigned to the training condition which played the AUC-DID. 20 subjects played the AUC-DID without biofeedback or arousal game influence. The effect of a one-shot training condition on the DE in the TIG appears to be only marginal (p=.147). Also game performance in the AUC-DIAG does not improve for the training group (m=33.88 correct decisions versus m=34.29, p=.810). Subjects receiving the ER training performed worse in the TIG, but also had increased phasic physiological reactions to gains and losses in the AUC-DIAG. Results in the performance in the AUC-DIAG hardly correlate with the Disposition Effect measured by the TIG (Spearman r=-.124; p=.40).

As mentioned before, subjects expose a physiological drop in HR directly after the experience of an emotion, where negative emotions typically result in a larger drop than a positive emotion. Often this emotional “imbalance” is accounted for biased emotional decision-making (cf. Bradley et al., 2008). Therefore, we employ the difference in subjects’ physiological reactions to gains and losses as a parameter for functioning emotional processing. For each subject an individual “imbalance” parameter is determined through the average emotional reactions to gains and losses in the AUC-DIAG. The parameter is significant with subjects’ individual game performance in the AUC-DIAG. However, we cannot successfully employ this parameter for a prediction of subjects’ exposedness to the disposition effect.

4.1.6.3 Implications for xDelia learning pathway design

In conclusion the results provide evidence that a single one shot intervention is not fruitful in improving subjects’ skills on effective emotion regulation. However, the results provide an indication that subjects’ emotional constitution, even between unrelated tasks, is a driver of the DE and that the AUC-DIAG elicits emotions that are relevant for decision making. With the results provided so far
we only mapped the physiological responses to gains and losses to the DE. Mapping physiological responses to regret on exposedness in the TIG were not successful.

The increase in disposition effect following an emotion regulation training intervention is puzzling but may be a short term effect due to exhaustion of regulative capacities. We speculate that whilst the long term learning effect of the game may be to improve self-regulation capabilities, the short term effect may be to temporarily exhaust self-monitoring capacity increasing reliance on system 1 emotion mediated decision paths.

4.1.7 S-M6: Transfer of learned emotion regulation skills from Aiming Game to Auction Game DID and Two Index Game (FZI-IPE)

Can the learned skills of emotion regulation be transferred from one task to another? The Space Investor game demonstrates the importance of emotional regulation skills in a more abstract manner whereas AUC-DID is more readily related to real financial decisions. Questions to be answered by this study are: (1) Can emotion regulation be trained and improved with the (AUC-DID) or a combination of the SI and AUC-DID (2) Can players which transfer acquired emotion regulation skills from one task (e.g. SI, AUC-DID) to another financial task (TIG)?

In Year 2, discussions between partners within xDelia lead to FZI deciding to collaborate with EUR (S-M7) to conduct a longitudinal study in parallel. The research aim remained the same but each institute then run a similar experiment examining the AUC-DID (FZI) and the SI (EUR) game. By dividing the workload a longitudinal study design could be carried out to examine the effects of repeated training of emotion regulation strategies (i.e., via biofeedback or mindfulness). By keeping studies consistent it gave EUR and FZI the option to compare analyses and even combine the data.

Four different types of training interventions (combinations of training interventions) were carried out at FZI.

- AUC-DID
- AUC-DID followed by SI
- SI followed by AUC-DID
- Control group (AUC without training (i.e. arousal was not permitted to have an influence on game difficulty and a measure of their arousal was not indicated on the screen)

Subjects received, depending on their assignment, three week training with the goal to determine which intervention was most promising for providing longitudinal training of emotion regulation. After each training session subjects played the TIG and the AUC-DIAG to assess changes in physiological reactions. Finally, in week 3, all subjects additionally conducted the Weber task (Weber & Welfens, 2008), as another measure for subjects’ exposedness to the disposition effect. The study procedure was chosen in close collaboration with S-M7 carried out at EUR.

4.1.7.1 Sample and methods

Altogether 108 subjects participated in the 3-weeks course with 27 subjects within each treatment. All subjects had no prior experience of undertaking emotion regulation training and most of them did not have an academic background in economics. Participants’ average age was 22.14 [18-27] years; 27 subjects were female.
All participants received, among other questionnaires, the Risk Aversion Questionnaire (Holt & Laury, 2002) and the trait ERQ (Gross, 1998) 1 week before the start of the study and 3 weeks after the study had finished. Moreover, subjects’ immediate emotion regulation state was questioned in each week by a state ERQ administered directly after each learning intervention was finished. Throughout the three week course, a baseline recording of HR was assessed before using the AUC-DID training tool. This was subsequently used for physiological analysis and for detecting development of HRV. To mitigate factors influencing HRV subjects from all treatment conditions for the duration of the course participated on the same day of the week and at the same time of the day.

Participants were notified that they would not receive their financial reward at the end if they missed their appointments for session 2 and/or 3.

4.1.7.2 Findings

4.1.7.2.1 Participant perceptions of the learning value of the intervention

4.1.7.2.1.1 via ERQ

The assessed ERQ for week 1 and week 3 is highly reliable for each subject. We can state that the overall learning intervention was successful in terms of subjects’ self-perception via this questionnaire. While suppression tendencies decreased significantly ($\Delta = -.251, N = 105, t(104)=2.523, p=.013$) over the course of three weeks, reappraisal tendencies tend to increase ($\Delta =+.184, N = 105, t(104)=1.568, p=.12$). This pattern is observable for both the treatment and the control group. When we make use of the state ERQ, which subjects received directly after the learning intervention in each week, the reappraisal score increased with each week for subjects that received emotion regulation training and drops for those that did not receive any emotion regulation training. In contrast, emotion regulation training decreases suppression tendencies with each week, while suppression tendencies rose on average for those included in the control group. In conclusion we can state the ER training increases subjects’ reappraisal tendencies, while suppression tendencies tend to decrease, as measured by the ERQ.

![Figure 4-9: Mean changes from in week 2 and week 3 (compared to week 1) in state Reappraisal scores and state Suppression scores between subjects that received training and subjects that did not receive any.]

4.1.7.2.1.2 via Heart Rate Variability

HRV (especially high frequency (HF) and the ratio between high-frequency / low frequency (LF/HF-ratio) reflect sympathetic and parasympathetic activity and are frequently employed as a
proxy of effective emotion regulation. High levels of HF and low levels of LF/HF-ratio indicate good emotional processing.

While resting HF-HRV\textsuperscript{12} remains unchanged from week 1 to 3 (0.120 vs. 0.119, N= 47, t(45)=-0.52, p=.958) for the control group, subjects that received ER training had a significantly higher HF-HRV in week 3 compared to week 1 (0.09 vs. 0.12, N= 138, t(136)=-3.105, p=.002).

In line with this we also observed that the LF/HF-ratio remains the same for the control group, whereas the ratio decreases significantly for participants that received emotion regulation training (3.685 vs. 2.612, N= 138, t(136)=-2.369, p=.019). It is also worthwhile mentioning that the observed HRV pattern was similar for all three of the ER treatments. This provides additional support that the emotion regulation training in fact resulted in enhanced effective emotion regulation by the means of HRV.

![Graphs showing changes in HRV between weeks 1, 2, and 3 for control and training groups.](image)

Figure 4-10: Mean changes from in week 2 and week 3 (compared to week 1) in HRV between subjects that received training and subjects that did not receive any.

Therefore, our findings demonstrate that the Training Interventions used in this study were effective for Emotion Regulation Training.

4.1.7.2.2 Relationship of Economic Results and Emotion Regulation Training

4.1.7.2.2.1 Two Index Game (TIG)

There was no significant difference between the control group and ER training on an absolute measure of the DE in the TIG. However, the DE measured in the TIG in the first session (week 1) did significantly and positively correlate with the DE measured in the Weber task (r=.208, p=.033) (an established research measure of disposition effect). This demonstrates that the TIG is a reliable measure of the DE, as measured by other tasks. Moreover, subjects performances in the TIG in week 1 and 2 (r=.608, p<.001) and 1 and 3 (r=.576, p<.001), and 2 and 3 (r=.688, p<.001) were correlated indicating that the DE in the TIG cannot be easily overcome by simply “learning” the game.

4.1.7.2.2.2 Auction Game Emotion Regulation

As expected, subjects become better playing the AUC-DID with each week they play it (decision performance i.e. percentage of correct decisions increases from week 1 (67%) to week 2 (75%) to

\textsuperscript{12} Resting HF-HRV effectively provides a measure of regulatory capacity.
week 3 (79%). There is no significant difference in subject’s improvement in decision performance between the treatments. However, when comparing subjects’ game performance in the 3 weeks it becomes evident that the training intervention was indeed successful. Subjects assigned to the Intervention Group with emotion regulation training had significantly lower arousal during the game and had a significantly higher variance on the arousal signal. In line with study S-M4, this indicates that the ER training and visual feedback helps both to keep arousal low, but also to regulate high levels of arousal down. Subjects in the ER training group, on average, also answered more decisions correctly on average. In line with our results from S-M4 – we find that high levels of suppression increase decision performance. Reappraisal, however, had no major influence on game performance.

4.1.7.2.2.3 Auction Game Diagnostic

Subjects become better with each week they play the AUC-DIAG. The groups with the learning interventions AUC-DID or AUC-DID and SI improved marginally faster in their performance compared to the control group. The correlation between the AUC-DIAG and the TIG appears to be week.

Each participant participated in a sequence of 50 decisions where one experiences joy, frustration and regret. From previous research we know that each of these emotions is reflected in a subsequent drop in Heart Rate. The amplitude of this drop is used for determination of an emotional imbalance in the subject, which could expose false decision behaviour.

Figure 4-11 depicts mean physiological responses to joy versus frustration and regret versus pride.

![Figure 4-11: S-M5: HR drop subsequently after the result is revealed after a positive versus a negative outcome (left) on a time axis (in seconds). As can be seen, the experience of a loss (incorrect decision) results in a stronger initial parasympathetic reaction (and hence heart drop), compared to the event of a gain (correct decision).](image)

We observe that physiological responses to losses (i.e. frustration) are correlated with game performance in the AUC-DIAG. However, in week 3 only joy remains significantly correlated with game performance. Interestingly, physiological reactions to gains are correlated with performance in the TIG in week 1 and physiological reactions to losses are correlated to performance in week 3. This provides more evidence that the TIG and the herein measures DE is dependent on emotions.

We cannot report any significance in task performance depending on the emotion characteristic regret or pride. While there is a significant development in physiological reactions to gains and foregone gains between the weeks, the influence of ER training remains weak.
4.1.7.2.2.4 Final performance in the Weber task after week 3
No direct effects were observed from the ER-Training on subjects’ performance in the Weber task. However, we observe that subjects high on reappraisal (measured by the ERQ in week 3) sell both winning assets and losing assets more readily than those low on reappraisal. This effect is consistent with another study conducted at FZI on the disposition effect and with studies S-M1&2.

4.1.7.3 Implications for xDelia learning pathway design
In conclusion the findings from this study show that the Learning Interventions tested are both successful in increasing subjects’ emotion regulation tendencies measured by the emotion regulation questionnaire as well as by the psycho-physiological mean of heart rate. However, the differences between the treatment conditions (AUC-ER, vs. AUC-DID + SI vs. SI + AUC-ER) are small, making it hard to judge the incremental contribution of the different interventions.

In this study the impact of training based improvements on tasks of economic decision-making such as the TIG, AUC-DIAG and the Weber task appear to be weak. So, while successfully boosting subjects level of emotion regulation, it remains unclear whether increased emotion regulation actually improves financial decision performance?

4.1.8 S-M7: Mindfulness and aiming game with the Two Index Game. Influencing the disposition effect: how effective are xDelia’s learning intervention tools? (EUR)
In Year 2, discussions between partners within xDelia lead to EUR deciding to collaborate with FZI (S-M6) to conduct a longitudinal study in parallel. The research aim remained the same but instead of conducting a series of short single-shot experiments with limited statistical power each institute would run a similar experiment either using the AUC (FZI) or the SI (EUR) game. By dividing the workload a longitudinal study design could be carried out to examine the effects of repeated training of emotion regulation strategies (i.e., via biofeedback or mindfulness). By keeping studies consistent, EUR and FZI have the option to compare analyses and even combine the data.

In previous studies (MF1, MF2), we established that short single-shot exercises on emotion regulation (short meditation) affected decision-making processes in students. However, it remained to be established whether these effects could be strengthened by practising these exercises over a longer period of time or in repeated sessions. In the present study, we aimed at investigating whether repetition of the same exercise over a limited period of time promoted learning effects of emotion regulation skills. The goal was to examine the learning effects of emotion regulation strategies (via biofeedback provided in the SI game, via a mindfulness exercise, or via a combination of both) over a three weeks period. The second aim was to determine whether a transfer of these skills occurred to affect the disposition effect (as measured in TIG and the Weber stock market task).

4.1.8.1 Sample and methods
Business Administration (n =112) students were recruited and randomly assigned to the following groups: Control (n = 28), Mindfulness (n = 28), Space Investor (n = 28), and Mindfulness x Space Investor (n = 28). In this study, 50% of the participants were female (n = 56) and the mean age of participants was 21 years (SD = 2.33).

Participants were recruited via the laboratory’s online bulletin board via advertisements. Upon arrival at the lab, participants received a short introduction to the experiment and completed informed
consent form. Next, participants completed a survey containing indices on trait emotion regulation (ERQ), trait mindfulness (Kentucky Inventory of Mindfulness Skills; KIMS), state mindfulness (Toronto Mindfulness Scale), and attention to bodily symptoms (Body Vigilance Scale; BVS). Next, participants were randomly assigned to one of the four following interventions.

- The Control group played Space Investor (SI) without biofeedback feeding into the game (hence, playing only a shooting game).
- The Space Investor group used SI(ER), which is a biofeedback game that helps players deal with their emotional arousal.
- The Mindfulness group received a mindfulness exercise, which was a meditation exercise that proved effective in inducing state mindfulness in previous studies (i.e., MF1 and MF2).
- The Mindfulness x Space Investor group received the mindfulness exercise each week, followed by SI(ER).

Then at the end of the first session, before being dismissed, participants were informed that they would not receive their financial reward for taking part in the study if they failed to attend their appointments for session 2 and/or 3.

All participants came into the lab once a week for three weeks in a row on the same day and at the same time of the day, where it was possible. Each session started with the weekly emotion regulation exercise (i.e., Space Investor, Mindfulness or Mindfulness x SI), followed by the independent variable: the TIG to measure their disposition effect. In the last week, we added another independent variable, the Weber stock market task to measure the disposition effect (Weber & Welfens, 2008).

4.1.8.2 Findings

Overall, students understood and complied with the study rules very well (the drop-out rate was less than 5%). Over a course of three weeks, students performed their emotion regulation training in the laboratory and reported afterwards that they tried to do this and play the games as best as they could.

4.1.8.2.1 Emotion regulation

As expected, trait emotion regulation (indexed by ERQ) correlated well between the first administration and the last administration of the ERQ-scales on Reappraisal ($r = .41; p < .01$) and Suppression ($r = .43; p < .01$). A MANOVA with within-subjects variable trait Reappraisal (before and after three weeks) showed that there were no differences in reappraisal scores between groups before and after the study ($F(3,108) = .65; p = .59$). Main factor Reappraisal indicated that all participants demonstrated a higher reappraisal score after the three weeks ($F(1, 108) = 1539.59; p < .01$). Although correlations may appear moderate for a stable, trait variable, this can be taken as evidence that changes in these variables occurred over the three weeks’ time period due to the emotion regulation exercises. To investigate this more closely for each week, we examined the state emotion regulation scores that were administered each week. First, we conducted a MANOVA with within-subjects variable Time (state RA in session 1, 2, and 3), and between-subjects variable Group (control, Space Investor, Mindfulness, mindfulness x Space Investor). The interaction term RA x Group was significant ($F(6, 216) = 2.2; p = .04$), indicating that the pattern of development in RA scores differed between groups. Figure 4-12 shows that this effect was mainly attributed to a significant gradual increase in state RA from session 1 to session 2. Follow-up contrasts revealed that, overall, a slight but significant decline was observed from session 2 to session 3 ($F(1, 108) = 4.41; p = .04$). The mindfulness group is the only group to show a consistent decline in state reappraisal. This is somewhat surprising, as mindfulness is associated with enhanced emotion regulation. It may reflect that in a mindfulness course, other kinds of exercises target reappraisal specifically compared to the 10 minute meditation exercise we asked participants to do. Perhaps, this specific meditation was effective in increasing attention levels or monitoring/noticing feelings but was less suitable for improving reappraisal specifically.
Then, we repeated the same analyses for trait suppression and state suppression, but found no significant effects (all p's > .20).

![Graph showing mean state reappraisal per measurement session](image)

**Figure 4-12: S-M7: Change in state ERQ-reappraisal over time and per training group.**

Note: C = control, SI = Space Investor, MF = Mindfulness, MF x SI = Mindfulness x Space Investor

### 4.1.8.2.2 Mindfulness trait and state

To examine whether trait mindfulness was affected by the weekly exercises, correlations were calculated between the scores on trait mindfulness (Kentucky inventory of Mindfulness Skills (KIMS)) at the start of the study and at the end of the study. This showed that trait mindfulness was a very stable factor (r = .75; p < .01).

To detect state mindfulness differences, we used the Toronto Mindfulness Scale (TMS). In our previous studies (MF1, MF2), the Decentering subscale of the TMS was most sensitive to changes in mindfulness states, so we explored that subscale here. We examined the scores on state mindfulness using MANOVA with Decentering (in session 1, 2, 3) as within-subject variable and Group (control, Space Investor, mindfulness, mindfulness x Space Investor) as between-subjects variable. As Figure 4-13 shows, scores on state mindfulness were consistently higher in the mindfulness and mindfulness x Space Investor groups compared to the control and SI groups. (F (1, 108) = 6.96; p < .01), indicating a successful mindfulness training.
4.1.8.2.3 Space Investor

First, we explored whether SI managed to increase arousal levels between the experimental and control participants. Univariate analyses revealed that compared to a control group, mean overall self-reported arousal (as defined by the xAffect software which feeds into the game) was significantly higher in both the Space Investor and mindfulness x Space Investor group ($F(1, 81) = 35.32; p < .01$).

Next, we examined whether the scores on paying attention to bodily cues (Body Vigilance Scale) varied during the three weeks. SI is a game where it is important to pay attention to emotional arousal, so we hypothesized that after playing SI, the use of bodily monitoring would have increased. We performed a MANOVA with within-subject variable Body Vigilance (in session 1, 2, 3) and between-subjects variable Group (control, Space Investor, mindfulness, mindfulness x Space Investor).

A significant interaction term revealed that there were group differences in body vigilance ($F(6, 216) = 2.52; p = .02$). Compared to the control group, all treatment groups scored equally and consistently higher on body vigilance ($F(1, 108) = 23.01; p < .01$) which supports the idea that training people with SI (and/or MI) helps them focus on their bodily signals.

4.1.8.2.4 Disposition effect

To examine the effect of each condition on the disposition effect (as indexed by the Weber task), we conducted a univariate ANOVA with independent variable Group (Control, Space Investor, mindfulness, mindfulness x Space Investor). No significant differences were observed between groups ($F(3, 110) = .76; p = .52$). Thus, participating in a weekly training in emotion regulation skills did not have effects on the size of the disposition effect in the Weber stock market task.
To examine participants’ disposition effect (as indexed by TIG), we first applied a log transformation to the disposition effect values. Next, to examine whether specific emotion regulation techniques affected their disposition effect, we performed a MANOVA with within subject variable DE (disposition effect score in session 1, 2, and 3) and between subject variable Group (control, Space Investor, mindfulness, mindfulness x Space Investor). Interaction term DE x Group was not significant ($F(6, 196) = 68; p = .67$), so the four groups did not differ with respect to the weekly disposition effect measurements. Instead, as Figure 4-14 shows, a gradual increase in the disposition effect over the three weeks was observed for all groups ($F(2, 196) = 7.8; p < .01$).

The disposition effect score in the Weber stock market task and in the TIG were significantly correlated ($r = .20; p < .05$).

The findings also demonstrated that all interventions impacted financial decision-making (i.e., higher DE in TIG). One explanation may be that the time interval between sessions (1 week) is too long and that the frequency of emotion regulation exercises needs to be higher to ensure that the player has internalised the emotion regulation strategy optimally.

When not having internalised the strategies sufficiently, one may wonder if actively trying to apply two new, cognitively demanding tasks (1. A continuous emotion regulation strategy, and 2. Trading) does not draw on cognitive resources too much, thereby depleting regulatory capacity in the short term and impacting financial decision-making accordingly. It could also be that it takes more than three exercises to establish more robust effects. It is not unlikely that more profound and intensive training courses can establish stronger changes in trait emotion regulation or trait mindfulness. When the individual has internalised these strategies, it is likely that results may vary as individuals do not consciously have to devote costly cognitive resources to applying this strategy. Yet, the control group also demonstrated an increase over weeks, which suggests that there is a learning effect to playing the TIG, which may cause players to start out a little cautiously in the first play-sessions. Thus, players may start trading more frequently after mastering the game basics, such that the probability that the disposition effect shows up increases.
Other data, including more behavioural patterns in the TIG, have yet to be analysed.

4.1.8.3 Implications for xDelia learning pathway design

The students showed changes in trait and state emotion regulation and in state (but not trait) mindfulness over the different emotion regulation training sessions compared to the control group. Body vigilance increased in all treatment groups compared to the control group as result of the emotion regulation training sessions, and the use of maladaptive emotion regulation strategies (suppression) declined across all participants. This indicates that the emotion regulation exercises were effective.

More importantly, this study on the repeated practise of the emotion regulation exercises indicated that a longitudinal training (e.g., a training of several weeks) is a valid approach as this weekly training session with simple exercises was already associated with an increase in trait emotion regulation (RA) (reappraisal), interoception (body vigilance) and state mindfulness before and after the study. This implies that, for example, the xDelia Mindful Trading Training where people log in daily to find a new exercise could be a helpful training programme to teach skills on emotion regulation and mindfulness. The present data are in line with this idea as the repeated mindfulness exercises indeed resulted in an increase in body vigilance (interoception) and state mindfulness (TMS). Such a daily exercising approach could potentially also be implemented for SI.

Interestingly, although both SI and mindfulness proved associated with emotion regulation, the effects of each group separately was for the most part equal or at least not much weaker than the combined mindfulness x Space Investor group. This suggests that combining the two trainings is not a necessary prerequisite for a training programme which aims at increasing emotion regulation skills.

The present data showed that although the treatment conditions were effective in increasing state mindfulness and emotion regulation skills, no effects were observed on the disposition effect in the Weber task. In contrast, a gradual increase in the disposition effect was noted for all participants on the TIG. It should be noted however, that the calculation of the disposition effect is not exactly the same in both tasks. It remains to be established how to interpret the gradual increase. It could represent a learning effect of playing the game, for example, that novice players will start more carefully than experienced players. This remains to be established.

The data indicated that a longitudinal approach could facilitate more robust changes on trait indices for long-term effects on emotion regulation, whereas in S-M7, only short-term effects on state variables were to be expected. The present data suggest that using a longitudinal approach for teaching emotion regulation skills is a good approach (according to the data, either mindfulness training or a training designed around biofeedback games, for example).

4.1.9 S-M8: Investor trials of the effect of mindfulness training, Auction Game and Space Investor Game training (OU, Saxo, EUR)

Due to changes in Saxo’s ability to make available access to a suitable numbers of clients in day trading centres we recruited cohorts of practicing traders (who invest on their own account) via trade shows (The World Money Show, Trader Expo and FOREX) and via an online trading forum. While this had the advantage of extending the diversity of investors we used in our evaluation study, it did
mean that we were unable to examine relationships to their real world trading data. Rather we used performance in the Two Index Game (TIG) as a proxy for trading behaviour.

This study combined an initial diagnostic phase with the biofeedback and mindfulness elements of the learning pathway and built on two pilot studies carried out at the Trader Expo and FOREX\textsuperscript{13} trade shows.

4.1.9.1 Aim of study

The aim of the study was to appraise the impact of playing the sensor games and mindfulness.

Research questions:

- Do the sensor games produce improvements in interoception?
- Does the combination of sensor games and mindfulness training produce a reduction in disposition effect?
- Do the target audience find these training interventions engaging and a useful basis for reflective learning?

4.1.9.2 Sample and methods

Participants (n = 58) were recruited from a number of different sources such as online discussion forums (e.g. via a Linked-In), and trade shows (such as FOREX London Investor Show and Trader Expo). This allowed us to draw on a pool of participants who, unlike our student samples, had actual trading experience.

4.1.9.3 Independent variables

Two variables were manipulated during the course of the study. The first was participants' engagement with the sensor games: the Space Investor game (SI) and the Auction Game (AUC). Both of these games can use ECG input from a Movisens sensor to calculate heart rate (HR) (via the xAffect system). HR is used in the game as a measure of arousal (calibrated against baseline HR). For each game the participant records a baseline period (typically 5 minutes) where they stare at a fixation point. This gives their base HR. During gameplay, if they become aroused or stressed, their HR will rise and this fall in HR (in relation to baseline) is fed back into the games making it harder (aiming becomes less accurate and targets move faster for greater arousal). The degree of arousal that the heart rate sensor is measuring is displayed graphically on the screen. The result of this is that if the participants playing the games become aroused then the games become more difficult. The games serve, therefore, as a biofeedback tool helping the individual to learn to manage their state of arousal and to pay greater attention to their internal emotional/physiological state.

This form of gameplay where the participant's HR fed back into the game was used in the ‘treatment’ condition.

The control condition is similar in that the heart rate data are recorded but not fed back into the game, nor displayed on the screen.

The second independent variable was training in mindfulness. In this instance a programme of mindfulness training, drawing on the extant literature, was used immediately after the laboratory session for the ‘treatment’ condition and then, using a waiting list control design, for the control

\textsuperscript{13} These are trade shows for private traders and combine talks and workshops on trading with exhibitions by companies who provide services to such traders.
group 2 weeks after the laboratory session. This training programme combined initial training instructions with four emails spread over a two week period. These emails provided additional exercises for the participants to help them further develop their mindfulness skills. To aid the practical application of mindfulness participants in the treatment group were provided with a trading diary which contained a number of questions about their mindfulness training, as well as asking for them to reflect on their trading. Participants in the control group received a diary asking for them to reflect on their trading.

In addition to the manipulated variables, data were also collected, using questionnaires, on a number of individual proclivities, such as emotional regulation and risk seeking, alongside trading behaviour and demographic factors. In some instances the questionnaires also served as a dependent variable (see below).

### 4.1.9.4 Dependent variables

Three different dependent variables were used. The first was the TIG which is an effective measure of the disposition effect, as well measuring a number of other trading related parameters. The TIG was administered at the very start of the laboratory session and then again, after the sensor games had been played. As the TIG is also available in a web version it was used as the dependent variable in the post-laboratory phase of the study. Participants were asked to play the TIG one week after the laboratory session and then again one week later (i.e. two weeks after the laboratory session). This enabled us to consider both the effect of sensor games on the TIG disposition effect and the subsequent effect of mindfulness training.

The second dependent variable was a measure of interoception, where participants were asked to tap out the rhythm of 35 heartbeats (x3) by clicking a computer mouse. Interoception is the individual’s ability to perceive their body’s own internal state and is related to emotion regulation capability; more effective emotion regulation is linked to greater self-awareness of emotional state (Barrett, Mesquita, Ochsner, & Gross, 2007). This ability, on the part of the individual, to ‘feel’ the degree of arousal that they are experiencing, is related to both the degree of arousal being experienced and the individual’s ability to perceive this arousal.

Our sensor games are designed to improve emotion regulation by offering the opportunity to practice regulating arousal levels. One important path by which this should operate is by improving interoception.

Thus interoception was measured before and after the sensor games to see if playing the games improved the individual’s ability to perceive their internal bodily state.

The third set of measures consisted of a questionnaire measuring user experience which was administered both at the beginning of the laboratory session and at the end of the study, two weeks after the laboratory session. This also included measures of mindfulness and emotion regulation strategy (the ERQ).

### 4.1.9.4.1 Study design

Participants were randomly assigned into one of two groups. The control group played the sensor games (SI and AUC) without the biofeedback either feeding into gameplay or being displayed on the screen. They also waited to receive their mindfulness training until after the study ‘proper’ had finished. The treatment group played the games with biofeedback and received mindfulness training over two weeks following the laboratory session.
4.1.9.4.2 Study implementation

Participants were recruited, as described above. They attended the Open University's Camden offices and prior to arrival were assigned to one of two groups; a 'treatment' or a 'control' group (waiting list). On arrival the participants had the purpose of the study explained to them.
After the initial explanation of the experiment the participants were presented with a volunteer information sheet and a consent form. Both of these documents, and the experiment itself, were approved by the Open University’s Research Ethics Committee. Participants then filled in a questionnaire which contained instruments to measure attributes such as habitual emotional regulation strategies, risk seeking, trading behaviour and demographic factors.

The participants were then fitted with the sensor belt for the heart rate monitor and given the tutorial for the TIG. They played the TIG as a baseline measure of the disposition effect. Participants then performed the interoception measure, again as a baseline against which any effects could be measured. Participants then played the sensor games, according to which group they were allotted to. This was then followed by a second interoception measure and a second playing of the TIG. Participants filled in a second questionnaire which solicited their opinions on the games. They were also interviewed about their opinions on the games and the overall approach. Finally the participants in the treatment group were given mindfulness training and the trading diary which contained both pages to record their trading and also to record their mindfulness activities. Participants in the control group were simply given their trading diary after they had completed the survey.

In the two weeks after the laboratory session the participants in the treatment group were given four emails containing further mindfulness training instructions. They were also invited to play the TIG one and two weeks after the visit to the laboratory. Participants in the control group were simply asked to play the TIG one and two weeks after the laboratory visit. After the second week (i.e. at the end of the study proper) they were then provided with the mindfulness instructions and training that the treatment group were given to ensure benefits for the control group.

4.1.9.5 Sample dropout

58 participants took part in the first phase of our study and they were allocated randomly to control (29) and treatment (29) groups. While we achieved full compliance from participants in the initial stage of the study, over the course of the two week follow up 19 failed to replay the TIG. Of the remaining participants several played the game multiple times (it turns out that investors are rather less compliant than students). In consequence, due to the low number of subjects, the power of our test of effects following the mindfulness intervention is reduced and results should be treated with caution.

4.1.9.6 Findings

4.1.9.6.1 Interoception

We carried out a repeated measures analysis of interoception comparing interoception before and after gameplay. (We ignored scores from the first measure since many participants initially found the task difficult to understand).

We found a significant improvement in interoception from before start of gameplay to after gameplay \( (F \text{ (time)} = 6.44, \text{ sig}=0.014). \) However we found no significant difference between treatment and control group in this change \( (F \text{ (time*condition)} =1.93, \text{ sig}=0.171) . \)

There are two plausible explanations: a) there was a difference between groups but the sample was too small to show a significant difference; b) the games do induce an improvement in interoception, but it is not dependent on arousal feedback in the game. This is plausible since the control group were also instructed that they should try and keep their arousal as low as possible. Although game difficulty for them was not dependent on managing arousal they will have been making efforts to manage and notice their arousal levels.
It is also possible that the improvement in interoception is a learning effect from taking the interoception several times in succession. This, however, seems unlikely since we asked participants to complete the interoception measure multiple times before and after gameplay and changes between interoception scores in multiple trials in this study show small and inconsistent changes whilst there is a larger and statistically significant change between the measures taken just before and just after the games.

### 4.1.9.6.2 Disposition effect

Due to the drop out in the second phase of the study we adapted our analysis in two ways. First we took the mean for each participant of disposition effect from the TIG in the follow up period. Second we used a multiple imputation approach to missing values as described for S-M1 and S-M2.

We carried out a repeated measures analysis of disposition effect over three time periods (before sensor games, after sensor games and during follow-up period, with trait emotion regulation approach (reappraisal and suppression) as covariates.

We found a significant relationship between suppression as an emotion regulation style and disposition effect (in all three periods) such that greater suppression was associated with greater disposition effect.

Results for changes in disposition effect by time and condition did not though achieve significance (at $p<.05$). However the direction of changes was interesting as may be seen in the figure below.
In the control group the disposition effect increases across the three periods. By contrast in the treatment group, disposition effect rises from before to after gameplay then falls to below initial levels across the follow-up period. This is consistent with a reduction in disposition effect consequent on the combination of gameplay and mindfulness training (from time 1 to time 3) but leaves the puzzle of why disposition effect rises in the immediate aftermath of playing the sensor based games.

One explanation is that self-regulation capability is like a muscle and can become exhausted (Muraven & Baumeister). Since both the sensor games and the TIG require active self-regulation, it is possible that the short term effect of playing the sensor games is to exhaust self-regulation capacity. Participants playing the TIG immediately after the sensor games then have depleted self-regulatory capabilities leaving them more prone to emotion driven biases.

We do though see an increase in disposition effect across multiple plays of the TIG in several studies. One possible explanation is suggested by comments made by several participants in follow-up interviews. A typical comment was “I wasn’t really being myself as a trader until I had played it a few times”. Further questioning suggested that in initial plays they were quite focussed on gameplay mechanics. It was only having mastered these that they felt free to focus on making as much money as possible. Thus it may be that the diagnostic value of the game is strongest after sufficient play to learn the game mechanics.
4.1.9.6.3 User experience

We asked all participants for feedback on their experience of engaging with the games.

Feedback was highly positive. 95% felt xDelia games could help them learn to manage their emotions and 100% of respondents agreed that their emotions impact how well they make financial decisions and that they could improve their trading decisions with better emotion management.

4.1.9.6.3.1  Space Investor Game

84% agreed the game was engaging to play and 82% that it was fun while only 15% felt it was to any extent boring. While fewer (54%) would recommend it to a fellow trader (23% disagreed) follow up discussion suggested this was mostly because they felt they had not yet had enough time to make a full judgement.

4.1.9.6.3.2  Auction Game

80% agreed the game was engaging to play and 77% that it was fun while only 7% felt it was to any extent boring. 78% would recommend it to a fellow trader (15% disagreed)

4.1.9.6.3.1  Additional findings

At the end of the training period, we did find higher mean scores on the state mindfulness inventory for the treatment group than the control group, although due to sample size this did not achieve significance at p<0.05. We did though find a significant difference in use of the reappraisal approach to emotion regulation among the treatment group following the training period (but no significant difference prior to this.)

We also find a significant inverse a significant inverse relationship between and mean disposition effect (across all TIG plays) and reappraisal scores. (r=.22, p<0.05) and a positive correlation between suppression scores and disposition effect (r=.36, p<0.05)

Finally we find good agreement between measures of TIG disposition effect taken at four different times (intraclass correlation coefficient = 0.74, p<.000)

4.1.9.7  Implications for xDelia learning pathway design

This study adds to evidence that the target audience find the xDelia learning games and mindfulness training engaging and likely to promote useful learning relevant to their trading behaviour.

The results on interoception suggest that not only do the games improve emotion regulation as found in previous studies but offer direct evidence that they may do so via improved attention to internal physiological states (interoception). The interoception tool developed as part of xAffect functionality and tested in this study may thus provide an important contribution to feedback and learning about emotion regulation.

The study provides further evidence that the TIG disposition effect measure has test retest validity given the moderately good agreement in DE scores across multiple plays.

While drop out among participants limited significance of findings, we also find the mindfulness training to improve both mindfulness and emotion regulation.

Evidence on disposition effects is weak and mixed. However, there is some indication that DE rises immediately with engagement in sensor based games and falls subsequently over a period of mindfulness training, in contrast to a steady rise for those not engaged in mindfulness training.
4.1.10 S-M9: Diary and reflection learning intervention

Information about private traders' views on trading diaries was collected during user experience interviews conducted after the studies Trader Expo and FOREX studies (see S-M8).

S-M9 then trialled the xDelia diary in paper form with a cohort of 58 traders as part of the 2-week longitudinal study S-M8(a). At the end of the initial stage of S-M8 in which private traders completed an emotion regulation questionnaire, played the games with sensors, the interoception measure and finally the TIG, they were given the diary and asked to record their emotional state as they traded. Pre-formatted pages were provided with check-boxes in which traders could record their emotions. At the end of the two weeks, participants were asked to return their completed diaries.

4.1.10.1 Sample and methods

A sample of 63 traders was recruited to take part in S-M8(c) and S-M9; of these 61 completed both studies. These participants all have experience of real world trading using an online platform and placing their own money at risk. Participants ranged in age from 18 to 77 and varied in trading experience from just starting out to 15 years, with the majority having between 1 and 7 years' experience of actively trading.

Participants were invited to attend the final stakeholder workshop, milestone M6.3, held at Saxo Bank on 15th May 2012. At this workshop, 34 participants took part in focus group discussions to gather their views on the diary experience and to give their input on what an online diary should look like.

After the stakeholder workshop and after all participants had received feedback on their results a follow-up web-survey was administered with a set of questions relating specifically to their experiences of using the diary.

4.1.10.2 Findings

4.1.10.2.1 Traders' and Investors' perceptions of the value of a Diary

User Experience Interviews: Prior to the S-M9 study, 90 private traders, who took part in Study S-M8 and associated pilot studies were asked about trading diaries. The consensus was that a trading diary was important, both for keeping a record of trades in order to track money made and lost, but also for recording a form of snapshot of the state of the market at the time of the trade. Some traders also described using the diary to record how they were feeling:

"So I will start with a bad day, or a good day, or felt good today, or some sort of expression on how the experience went just naturally seems to come mind first when I fill in my log. And I guess learning points are probably the next thing I’d write down and just obviously the accounting side, profit or loss, I’ll write that sort of thing down as well."

Many traders replied that they believed that a trading diary was a valuable tool, however, they didn’t currently keep one although they felt they ought to,

"I have occasionally. I know I should do but I don’t."

Several respondents made the point that they found a trading diary to be a most useful tool when they were losing money and so tended to lose the routine when their trades were doing well:
"I used to when I was losing. Since I'm winning I'm not keeping it. I don't know why. I should keep I think."

"it's whether you keep the routine and then when things start going ... if you're going through a period where things are going well and you tend to kind of let it slip a bit."

Post-S-M8 User Experience Web Survey: In the web survey (N=20) administered after the S-M8(c) study, most (70%) participants already kept a trading diary (35% for every trade, and 35% for 4 out of every 5 trades). 15% responded that they believed a diary was important but didn't keep one at present and 15% said that they didn't feel the need to keep a trading diary. When asked whether they felt that keeping a trading diary was good practice, 90% responded yes, 10% unsure.

When asked if they had used the xDelia Trading Diary, with its check boxes for recording emotional state before and after trading, 68% had done so, (16% for every trade, 5% for 4 in 5 trades, 16% for approx. 2 in 5 trades and 16% intermittently).

The majority of participants were positive about the value of recording emotions in the xDelia Trading diary, with 67% agreeing that keeping the diary helped them manage their emotions effectively during trading (22% undecided, 11% disagree, 0% strongly disagree). 67% felt that recording their emotional state before and after trading made them aware of their emotions (22% undecided, 11% disagree, 0% strongly disagree) and 50% felt that this awareness of their emotions during trading helped them make better trading decisions (44% unsure, 6% disagree). Part of the role of the xDelia Trading Diary is as a reflective mechanism, and 67% of participants reviewed their previous diary entries and felt that reviewing their previous diary entries provided insights into the role of their emotions in their trading decisions.

4.1.10.3 Implications for xDelia learning pathway design

The pre S-M8 interviews and the S-M9 survey responses confirmed that our diary-based approach to supporting reflection and emotional awareness during trading was seen as valid by our target stakeholders, and that the support for recording emotions provided in the pilot xDelia diary promoted emotion awareness which participants felt would lead to better trading decisions. The S-M9 focus group, held during the Stakeholder workshop in M39, allowed us to give our participants the bigger picture of the xDelia research on which these studies were based, and to gather their input on the design of the online diary tool.

All 58 participants in S-M8 were invited to the Saxo Bank final stakeholder workshop (M6.3), and 34 attended. Their input was gathered during a focus group. Jeff Lins of Saxo Bank demonstrated features of a planned automated diary system and solicited reactions to these plans and to their use a diary during the prior study.

Key points which emerged:

- Slightly over 50% currently keep a journal or trading diary.
- Recording emotions or feelings is important as it is easy to forget the next day.
- About 1/3 of attendees already use their journals to capture emotional aspects.
- About another 1/3 of attendees would find it useful to use an online trading diary to capture emotional state with some still unsure.
- Weekly review of diary entries is important as learning only happens if entries become a basis for reflection.
- Diary very useful for reflecting on bad trades.
- The ability to automatically capture trading data including entry and exit points and market prices over the period is very important. Some already do this by capturing screenshots but
this is an awkward way of achieving the goal and does not allow drilling down into other
data.

- Consensus was that previous day's entries should not be editable to avoid deluding yourself. However annotating entries after the event would be a useful function.
- The ability to track transactions against other events (timestamps) over a long period would be very useful.

An overwhelming majority of attendees were in favour (almost 100%) of our approach and were very positive about the idea of integrating sensor based physiological feedback into the diary.

An important issue that was emphasised by several participants was the value of automated capture of trading information and the facility for rapid capture of emotion state. This was they explained as a result of their experiences of starting and stopping diaries several times because the effort needed to record all relevant information became impossible to sustain.

4.1.11 S-HPa: Heuristic Evaluation and Playtesting of Auction Game

Work on the S-HP series of studies, was primarily carried out as part of Work Package 4 and fed into iterative improvements in the learning elements. They are reported here for completeness. Further information may be found in the Work Package 4 deliverable D14.

Development of the Auction Game (AUC) was followed by a heuristic evaluation. This evaluation aimed to qualitatively identify design errors and suggest improvements. The heuristic evaluation (Isbister & Schaffer, 2008) used on the AUC is part of a generic evaluation tool kit which is being developed and used on all prototypes produced at BTH developed within xDelia. The heuristics are divided into a set of categories for inspecting different aspects of the game prototype. Framework and evaluation results are presented below. During the evaluation, if two or more reviewers agreed on an issue, it was immediately noted.

Playtesting is valuable since it allows different players to subjectively analyse the game. This may reveal novel aspects that have not previously been discussed or evaluated by the development team. The study reported here was an individual case study using both quantitative and qualitative data. Two cases representing participant categories of trait emotion regulation were identified: suppressors and reappraisers of emotions. Subjects were separated in to these different groups when analysing data using results from the Emotion Regulation Questionnaire (ERQ) (Gross, 1998).

The heuristic evaluation was carried out by three evaluators having competence both in games and in usability. The evaluators evaluated the game separately and no collaboration was allowed. The list of the heuristics was distributed to the evaluators, containing criteria based upon which they described the violating issues. The task was to find the problems, while dealing with the problems was left to the development team. Following this, the evaluators met and constructed a joint list of issues; such that, if two evaluators had the same issue on their lists, the problem was transferred to the final list. All of the issues reported by a single evaluator were discussed and only transferred to the final list if the whole team agreed that the issue was in fact a problem. A report was prepared (DOC-AUC-5 “Results of heuristic evaluation”) describing the issues in more detail, accompanied by screenshots to clarify the issues. The evaluators together with the developer discussed possible solutions and the suggestions were fed into the Final specification of the Auction Game (DOC-AUC-3.1) document. After the Heuristic Evaluation was conducted and documented, the results were presented to the product owner, who in collaboration decided on initiation of the playtesting studies.
Before playing the game, the playtesting participants were fitted with the Movisens ekgMove ECG sensor and given a tutorial session. In order to objectively determine which game elements the players paid attention to, the game was played using a Tobii T60 eyetracker logging data on different Areas of Interest (AOI) and recording the whole gaming session on video. The purpose of using the eyetracking software was to assess the importance of different visual objects (AOIs) on the screen to a player. At the end of each game, each participant was given the ERQ (Gross & John, 2003) in order to identify suppression and reappraisal tendencies of individuals and a modified System Usability Scale (SUS) (Brooke, 1996; Nacke, Schild, & Niesenhaus, 2010) questionnaire measuring game usability. The SUS questionnaire contained 10 questions, the sum of which represented a composite measure of the overall usability of the game being studied. An interview session was also conducted where participants could openly discuss perceived game speed and difficulty, as well as visual cue elements and any other issues they wanted to note.

4.1.11.1 Sample and methods

A total of 6 students volunteered to participate in the Playtesting Evaluation. They were all students of BTH aged between 20 to 32 years of age; of which four were male and two were female. Participants reported varying previous gaming experience.

4.1.11.2 Findings

4.1.11.2.1 Heuristic evaluation

Heuristic evaluation pinpointed several important design issues and reinforced the robustness of the game. For more detail about heuristic evaluation please refer to the DOC-AUC-5 document (results of heuristic evaluation).

4.1.11.2.2 Playtesting evaluation

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

The game was successfully played up to the 8th level by two participants, both of whom were 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 that the AUC is a good as a tool for emotion regulation training. The rest of the participants got stuck on the first level struggling with the game mechanics and speed. This provided justification for making the beginning levels a bit easier so every participant can experience motivation in progress.

Eyetracker gaze data analysis clearly identified areas of player interest that perfectly correlated with visual cue elements presented on the screen. Identified elements were: estimation clouds (which change colour) in the centre of the screen; buy/sell buttons in the bottom part of the screen; the arousal meter indicator in the top right corner; the money pile in the bottom right corner of the screen.

Because of the fast pace of the 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 hand corner of the screen. We evaluated this claim on how informed the participants were about their arousal level by keeping track of it on the arousal meter indicator during rounds in the whole gaming session. A paired-samples t-test was conducted on the eyetracker data to evaluate the difference in number of gaze observations on a marked indicator arousal meter AOI compared to the number of rounds taken for each participant. There was a statistically significant difference found with the number of rounds (M=110.17, SD=95.26) to the number of arousal meter observations (M=16.33, SD=24.5, t(5)=2.94, p=.032). Thus we can say that participants paid little or no attention to the arousal meter.
indicator during the whole playing session. To rule out the fact that two advanced participants used the information on the arousal meter indicator to perform better in the game, a similar test was conducted comparing focus time on marked indicator arousal meter AOI for each participant and no significant difference was found. Participants reported that the reason for paying little or no attention to the arousal meter was due to the limited amount of time that was available given the fast pace at which decisions had to be made. Most of the participants reported that they were paying attention to their arousal indicated by the colour of the cloud estimations, especially when it turned red. This gave justification to concentrate on making the colour coded cloud estimations more distinct rather than focusing on further developing the arousal meter.

<table>
<thead>
<tr>
<th>Arousal level while making a decision (Groups)</th>
<th>Money earned</th>
<th>Time needed for decision making</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. dev.</td>
</tr>
<tr>
<td>1</td>
<td>0.63</td>
<td>2.95</td>
</tr>
<tr>
<td>2</td>
<td>.46</td>
<td>3.24</td>
</tr>
<tr>
<td>3</td>
<td>-0.47</td>
<td>2.82</td>
</tr>
<tr>
<td>4</td>
<td>0.67</td>
<td>3.29</td>
</tr>
<tr>
<td>5</td>
<td>-1.36</td>
<td>3.03</td>
</tr>
</tbody>
</table>

Table 4.3 – Money earned and Time needed to reach a decision at different arousal levels

A one-way, between-groups, ANOVA was conducted to explore the impact of arousal level on money earned in each round. The total number of 661 rounds played was divided into 5 groups according to their arousal level while making a decision (Group 1: [relaxed], Group 2: 2 … Group 5: [highly aroused]). There was a statistically significant difference at the p<.05 level in profit made in each round for the five arousal groups [F(4, 656)=3.566, p=.007]. The effect size, calculated using eta squared, was .02. Post-hoc comparison using a Turkey HSD test indicated that the mean score for Group 1 (M=.6328, SD=2.95) was significantly different from Group 5 (M=-1.369, SD=3.3). Other groups did not differ significantly. The same analysis has been conducted on the time needed to reach a decision in seconds and there was a statistically significant difference, with p<.05 [F(4, 656)=5.753, p=.000] between Group 5 (M=1.55, SD=.45) and the rest of the groups. The effect size, calculated using eta squared was .03. This gives strong evidence confirming that the design of the AUC rewards a player for achieving a target arousal level and increasing their earned profit, while at the same time presenting a hard challenge and punishment to a player in an undesirable high arousal emotional state.

4.1.11.3 Implications for xDelia learning pathway design

4.1.11.3.1 AUC as a tool for raising awareness and training of emotional regulation

Evidence shows that emotions have an in impact on decision making, especially in the field of finance. Thus it makes sense to develop a tool to help people to become aware of this as well as to help them regulate their emotions to achieve better financial decisions. A serious game emerged as an effective tool for the job where players receive instant emotional arousal biofeedback information on screen which they can interact with, and choose to regulate it.

During fast paced decisions, players concentrate on the information presented and do not have enough time to shift their attention away from the task at hand. As participants reported and the data showed, they paid little or no attention to the arousal meter indicator, while the arousal indicated by the cloud estimations was perceived and used, especially when the clouds turned red on high arousal
values. After discovering this, we concentrated on making the colour of the cloud estimations more distinct, since this is where players were focusing their attention.

In the interview session, it was found that novice participants perceived the game to be too fast/just bearable, while two successful participants perceived it as being manageable since they were practicing emotion regulation (relaxation) techniques. Successful participants reported that the game was quite overwhelming at times and that sometimes they needed to let a decision pass, thus losing profit, in order to relax and win in the next set of decisions. This gives evidence that the AUC is indeed overwhelming and capable of putting players in highly aroused state where they need to practice emotion regulation techniques to succeed in the game.

4.1.12 S-HPb: Heuristic Evaluation and Playtesting of Space Investor Game

The aim of this study was to implement a second version of the Aiming game (called the Space Investor game), through functionality and play testing while learning from many of the design flaws of the first version. This was to be done in close cooperation with the product owner (PO) - EUR – whose role it was to take critical decisions regarding development since they were also responsible for conducting the main study.

The implementation was conducted in an incremental setup so that there was always a functional game available for partners to use. When there was one month left it was demonstrated to the PO who got to have opinions on the game’s design and functionality. Thereafter, the product was finalized to fit the PO’s requirements that were reasonable due to implementation time. When the studies were postponed, more time could go into development and, with the help of the PO, the game was fine-tuned according to their requirements.

Playtesting was also performed before shipping the game to the partners (EUR and OU) for more in-depth studies.

4.1.12.1 Sample and methods

The play testing evaluation of the Space Investor (SI) game was done with five participants who all played the version used in the longitudinal study performed by xDelia-partners at Erasmus University. All participants used the Movisens EKG sensor with a 300 second baseline. Five levels, excluding tutorial and intermissions, with a length of up to 180 seconds were played. All players filled in the game experience questionnaire (GEQ) (Nacke, 2009; Nacke and Niesenhaus, 2010) and system usability scale (SUS) (Brooke, 1996).

4.1.12.2 Findings

The results showed a spread in difficulty (GEQ 13), enjoyment (GEQ 1, 41), immersion (GEQ 5, 10, 15, and 40), stress (GEQ 7), perceived skill (GEQ 17), boredom (GEQ 18), irritation (GEQ 24), challenge (GEQ 29), and time pressure (GEQ 37). These results suggest that the SI game might not be a game for everybody. While several participants showed positive results in the above mentioned categories, others did not, suggesting that another game might be more suitable for them.

Most participants did not appreciate the story of the game. According to the play testing, very few players felt that it contributed to the game experience at all (GEQ 3, 35).

While most of the play testers did not feel completely stimulated (GEQ 26) by the game, nor felt like they learned a lot (GEQ 8), the majority felt happy during the session (GEQ 6, 38). It is likely that these results are due to play testers’ expectations and gaming experience and might not transfer to the
investor setting. On a positive note, however, the more experienced gamers in this play testing sample did find the game to be aesthetically appealing (GEQ 14). Participants experienced the game as limited in terms of exploration possibilities (GEQ 21). The game is, however, intentionally very linear in order to control content and could have been made to appear less linear if this had been deemed necessary.

One of the most important aspects from the Aiming Game was the intuitive interface and gameplay mechanics. According to the SI play testing, players congruently felt that this aspect had been sustained, if not improved (SUS 1-10). All items of the SUS questionnaire consistently gave good usability results pointing towards usability being preserved between the Aiming Game and the SI game.

4.1.12.3 Implications for xDelia learning pathway design

Since the SUS questionnaire revealed that usability had been sustained since the Aiming Game, and that the SI is considerably longer (and easily scalable on top of that), partners were advised to use the SI game in order to create a better learning environment for emotion regulation training.

It was also concluded that more could also be done with regards to the story of the game and possibly getting participants to compete against each other in the training area by showing a score board.

4.1.13 S-HPc: Initial Usability and Playtesting of xHale Athlete

4.1.13.1 What we said we would do

S-HPc involved the usability and play testing of the different interventions in xDelia’s learning pathway (how the game is used and how the users experience the game). Laboratory studies were not definitely planned; there sole purpose was to provide input to the design process.

Both the usability and play testing of the mindfulness game (xHale Athlete) and the mobile phone application (xMedit-Mate) were performed as planned. However, it should be recognised that both the game and the application were planned and developed after submission of D9-2.3.3 and were therefore not part of the original study plan submitted in Year 2.

4.1.13.2 Sample and methods

4.1.13.2.1 xHale Athlete

The usability testing was done with four participants (1 female and 3 males aged 18-45). Two of the players were very inexperienced gamers, while two of the player would normally play computer games several hours a day. Only one of the players claimed to be experienced in the platform genre, which was confirmed by the final score of the game.

In the usability and play tests done the researcher acted as a test leader. The players sat alone with the researcher in a room and the only instruction they were given was that the game was about mindfulness. Before they started to play the game, they were asked to fill-in a questionnaire to collect demographical data and questions about how experienced they were at playing platform games and whether they knew about mindfulness. The player played the game while the researcher observed and took notes of how the player acted.
After the play session the players answered the Game Evaluation Questionnaire (GEQ) mainly focusing on their play experience, and the System Usability Scale (SUS) focusing on the usability of the game. After each play session the player and the researcher discussed the game. What was good and what could be improved? How did the player experience the game? What was easy and what was too difficult. The players were also asked if they got curious about mindfulness after playing the game. The players only played the game once.

4.1.13.2.xMedit-Mate

The xMedit-Mate was only usability tested because it is an application supporting the learning intervention rather than an intervention in itself. The test was conducted by five users. One of the five users had practiced mindfulness before. Neither of the other users knew anything about mindfulness.

The inexperienced users were briefed about mindfulness and what it is, before the test started. The researcher acted as a test leader and gave the user some scenarios to accomplish. The scenarios involved using the different parts of the application. For example ‘Set the mediation timer to ring between four and six in the afternoon’. The user was asked to talk aloud while using the application. The time for the user to complete the task was recorded and the test leader was observing what the user was doing and made notes about what was going on. After each scenario the user and the test leader discussed the user’s experience of the task. After all the scenarios were finished the user filled in the System Usability Scale and the test ended with a discussion between the researcher and the user about the usefulness of the application.

4.1.13.3 Findings

4.1.13.3.1 xHale Athletic

All of the users did find the game very easy to use and no one had questions about how to play the game. The only difficulty was the paced breathing exercise. All of the players found the game hard. The experienced user managed to finish all levels. To count the platforms at the same time as you need to keep track of ducks or moles was experienced as extremely difficult especially as the platforms are so many. The breathing and mindfulness were experienced to get out of focus while concentrating on several tasks at a time. All of the players experienced the levels with automatic scrolling as very stressful and it was hard for all of them to keep up.

One of the purposes of the game is to invoke curiosity about mindfulness and act as a portal to the mindfulness course. Three out of four players said they would like to know more about mindfulness after they have played the game.

4.1.13.3.2 xMedit-Mate

The users explored all elements of the xMedit-Mate: calibrating and performing a paced breathing task, using the meditation timer, using the meditation planner and the feedback menu. All the users had problems to understand how to make the calibration, in the breathing exercise, work. The test leader had to help out, but when they got the breathing exercise to work they found it quite good, except for the experienced user that did not think it added anything to the mindfulness training. The experienced user experienced it as the visual pattern (expanding and shrinking circle) draw the attention from the breathing itself.

All the users managed the meditation timer without trouble. As the users had performed the paced breathing just before this exercise they all practiced the breathing again. The experienced user closed the eyes while the other users stared at the progress bar shown at the screen.

The meditation planner was perceived as fairly easy to use. The experienced user thought that the randomized time periods were something that would not be used. The experienced user preferred to use the fixed time options.
The chart displaying the statistics of for how long and how many times the user had practiced mindfulness resulted in some issues of understanding about what the chart showed.

4.1.13.3.3 Implications for xDelia learning pathway design xHale Athletic

There were two major issues in the game, one concerning usability and functionality and a second one that might interfere with the intention of the game, that is mindfulness. First, the users experienced that it was difficult to understand how to use the key to make the circle shrink and expand in the breathing exercise. Due to this issue the breathing exercise has been redone and tested separately, and is easier to understand in the latest version.

The second issue concerned the task to count platforms while playing to introduce some interference in the task to notice flying objects. The players experienced that it was easy to lose track of how many platform the character had jumped on as there are so many platforms.

When it comes to how fun the game is there were different opinions. Two of the users thought it was quite fun. The most inexperienced player and the most experienced player found the game fun. It appears like the inexperienced user concentrated on jumping and getting coins and to succeed to do that was fun enough. Attesting to the validity of the game as sensitive tool to trait differences in mindfulness, the experienced user scored much better than the others and tried to take all of the coins which appeared to motivate him.

In the discussions there were suggestions of how to improve the game. For example, the two gamers wanted to have some power-ups, some enemies to avoid, other ways to collect coins e.g. look for them in barrels or other containers, to make the player produce coins by doing something like hitting the head in special platforms and have different kinds of platforms with special behaviours and attributes. After the discussions with the players a new feature (shoes with wings) was implemented. If the player collects the shoes the character moves faster which means that the player himself can choose how difficult the game should be, and if played repeatedly the game can stay challenging.

Studies in Work Package 2 during the final year have had a primary focus on evaluating the xDelia learning pathway and its elements, which we designed in Year 2. (A detailed account of the learning design may be found in the Year 2 deliverable D23-2.5: Pedagogic Framework, and in the end of project deliverable D18-2.4.2: Intervention package – specification and development).

4.1.13.3.4 Implications for xDelia learning pathway design xMedit-Mate

From the user test it was revealed that the calibration of the breathing pace was hard to understand. It resulted in a re-make of the calibration and better instructions. The experienced user also questioned if the circle in the breathing exercise added anything to the mindfulness experience or if it just draws attention from the intended activity. Future research should investigate how this feature is experienced in studies using the xDelia mindfulness training. Another issue is if the progress bar in the meditation timer influences the meditation positively or negatively.

The chart displaying the statistics of for how long and how many times the user has practiced mindfulness resulted in some issues of understanding what the chart showed. Therefore better descriptions of the four different charts were introduced in the last version. The outcome of the test also indicates that the usefulness of the statistics must be investigated further.
5 References

Greiner, B. 2004. An online recruitment system for economic experiments.


