The investigation of reward-based learning in obsessive-compulsive subclinical checkers

By

Camila Gomes Victorino

Thesis submitted to the University of Surrey in partial fulfilment of the requirements of the degree of Doctor of Philosophy

School of Psychology

Faculty of Health and Medical Sciences

University of Surrey

June 2017

© Camila Gomes Victorino 2017
Declaration of originality

This thesis and the work to which it refers are the results of my own efforts. Any ideas, data, images or text resulting from the work of others (whether published or unpublished) are fully identified as such within the work and attributed to their originator in the text, bibliography or in footnotes. This thesis has not been submitted in whole or in part for any other academic degree or professional qualification. I agree that the University has the right to submit my work to the plagiarism detection service TurnitinUK for originality checks. Whether or not drafts have been so-assessed, the University reserves the right to require an electronic version of the final document (as submitted) for assessment as above.

Camila Gomes Victorino

Signed:

Date: 12th June 2017
Abstract

People with Obsessive Compulsive Disorder (OCD) have deficits in decision making under ambiguity, also known as probabilistic reward-based learning (RBL). In these tasks participants are not aware about the target probabilities of each option and need to use trial-by-trial feedback to learn these probabilities during the task in order to improve their performance. An open question from the previous literature is why OCD participants present this specific deficit. The review of the limitations of previous RBL studies with OCD patients and the use of a new task paradigm allowed us to explore new directions in this research field. Firstly, previous studies had the following limitations: sometimes they did not control for the use of medication, for the presence of common comorbidities associated with OCD, or for differences between distinct OCD subtypes. In addition, previous tasks usually manipulated several factors at the same time, such as feedback direction (positive vs. negative feedback) and feedback magnitude (e.g. less vs. more positive feedback), for instance. Therefore, it is not possible to know which factor affected the RBL performance in OCD patients. Furthermore, previous studies did not consider the effect of symptom-related feedback on RBL. Secondly, task performance was only measured with accuracy, even though RBL involves the use of previous trials to make predictions for the current decision. Finally, while it is reported that participants do use distinct previous trials to perform these tasks, sequential effects were not previously investigated in OCD patients, neither for random-sequences nor for patterns sequences.

Based on this, the presented PhD thesis addressed some of these gaps in the literature. The investigation was restricted to subclinical checkers which is the most common OCD subtype. A binary decision making task was designed that allowed us to separately manipulate feedback direction and feedback magnitude. In addition, the study compared probabilistic RBL performances between checkers and non-checkers for random and pattern sequences. New analyses techniques were employed to investigate the use of previous trial information for the current decision making, e.g. win-shift, lose-shift, and cross-correlations.

Studies 2 and 3 examined the effect of feedback direction (positive vs. negative feedback) on decision making in subclinical checkers. Results showed that subclinical checkers were more biased towards exploitation when using negative feedback. Studies 4 and 5 examined the effect of the feedback magnitude for both the positive and the negative feedback direction. Results showed that subclinical checkers were able to change their bias
towards exploitation within the positive direction experiments depending on the feedback magnitude of the experiment. They were more biased towards exploitation in the positive task presenting a higher error magnitude. In addition, checkers were always biased towards exploitation within the negative direction experiments, irrespectively of the feedback magnitude. Study 6 examined the effect of the manipulation of feedback magnitude when a symptom related feedback was added as an increment of the negative feedback.

Results showed that subclinical checkers were not affected by the presence of the symptom-related feedback and, in terms of feedback magnitude, they continued to present a bias towards exploitation. The findings show that checkers were able to adapt towards exploitation in the positive experiments when higher feedback magnitudes are given. This strategic shift might be related to the higher error magnitude associated with the absence of reward, when the incorrect option was chosen. In contrast, checkers were not able to adapt their exploitation behaviour within the negative experiments when enhancing the feedback magnitude. In this way, this indicates that subclinical checkers present a deficit that bias their responses towards exploitation, when the magnitude associated with the error in the task surpasses a certain value of negative magnitude. This bias could reflect deficits regarding negative prediction errors in OCD or a hyperactivation of brain areas related with exploitation. Both explanations could be linked with a hyperactive dopaminergic system in OCD, so these results could encourage new research about the role of dopamine and prediction errors deficits associated with OCD. Additionally, heightened emotions and reward magnitudes might reduce treatment success because of the enhanced exploitation behaviour, so one crucial aspect of future therapies might be to carefully study the employment of stimuli with such higher error magnitudes.
Acknowledgements

This thesis would not be possible without the help of my family and friends. In particular, I would like to thank my beloved husband Jonas Guedes Ribeiro, who encouraged me during all these four years and was always present no matter what. The process of developing a PhD is not easy but even harder is to be distant from your family and country during these years. Actually, besides the PhD challenges, having to adapt to a different culture and language would not be possible without his presence in my life, always encouraging me to continue and giving me hope to pursue my dreams.

I would also like to thank my mother who inspired me to study and always supported me when I needed. My brother and sister were also always there, when I needed a good talk. Finally, I would like to thank my father who was always an example of intellectual curiosity for me. I hope he is seeing me from somewhere else and receive my thanks for his example in my life.

In addition, I could not forget to mention two people who were very important for me during this PhD. First of all, I would like to thank Dr. Marcus Vinicius Baldo, who always believed in me and helped me to get this PhD position at University of Surrey. Vinicius, as I call him, was a great supervisor and is a good friend who had the sensibility to understand I needed more than a simple academic support to follow my dreams. Thank you so much for your help and for believing in me.

Secondly, I would like to thank Dr. Ellen Seiss. Dr. Ellen Seiss was not a common supervisor, she was more than that. She was always present when I needed her, as a friend and as a supervisor. I would like to thank you for your support, for always inspiring me to become a better academic and for being there. I was very privileged to have a supervisor like you.

Finally, I would like to thank Dr. Annette Sterr. She was an excellent supervisor and was a good example of academic curiosity in my life. Additionally, I would like to send a special thanks to Dr. Johann Briffa who was a collaborator in my project and helped me to design the experimental sequences used in this thesis. And finally, I would like to thank all my friends who did not forget me during these four years abroad.

Thank you all for your support and to help me to remember I am loved.
# Table of contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>i</td>
</tr>
<tr>
<td>List of tables</td>
<td>vii</td>
</tr>
<tr>
<td>List of figures</td>
<td>xi</td>
</tr>
<tr>
<td>List of abbreviations</td>
<td>xii</td>
</tr>
<tr>
<td>List of appendices</td>
<td>xiv</td>
</tr>
<tr>
<td>1. Chapter 1. Literature review</td>
<td></td>
</tr>
<tr>
<td>1.1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2. Obsessive compulsive disorder</td>
<td>2</td>
</tr>
<tr>
<td>1.3. OCD models</td>
<td>9</td>
</tr>
<tr>
<td>1.4. Typical OCD treatments</td>
<td>14</td>
</tr>
<tr>
<td>1.5. Emotional and cognitive deficits in OCD</td>
<td>16</td>
</tr>
<tr>
<td>1.6. Reward based decision making and learning in OCD</td>
<td>19</td>
</tr>
<tr>
<td>1.7. Feedback direction and magnitude manipulations in RBL</td>
<td>28</td>
</tr>
<tr>
<td>1.8. First order and second order (sequential) effects in RBL</td>
<td>31</td>
</tr>
<tr>
<td>1.9. Summary and overview of experimental chapters</td>
<td>39</td>
</tr>
<tr>
<td>2. Chapter 2. General methods and data analysis</td>
<td>44</td>
</tr>
<tr>
<td>2.1. Introduction</td>
<td>44</td>
</tr>
<tr>
<td>2.2. Experimental and overall procedure</td>
<td>45</td>
</tr>
<tr>
<td>2.3. Participants payment</td>
<td>45</td>
</tr>
<tr>
<td>2.4. Part 1 - Online Screening Questionnaire</td>
<td>46</td>
</tr>
<tr>
<td>2.5. Part 2 – Laboratory-based Study</td>
<td>49</td>
</tr>
<tr>
<td>2.6. Data analysis of the laboratory-based part of the study</td>
<td>63</td>
</tr>
<tr>
<td>3. Chapter 3. Picture validation task</td>
<td>72</td>
</tr>
<tr>
<td>3.1. Introduction</td>
<td>72</td>
</tr>
<tr>
<td>3.2. Methods</td>
<td>73</td>
</tr>
<tr>
<td>3.3. Results (with discussion)</td>
<td>77</td>
</tr>
<tr>
<td>4. Chapter 4. Pilot experiments</td>
<td>87</td>
</tr>
<tr>
<td>4.1. Introduction</td>
<td>87</td>
</tr>
<tr>
<td>4.2. Methods</td>
<td>88</td>
</tr>
<tr>
<td>4.3. Piloting sequences (Results and discussion)</td>
<td>92</td>
</tr>
<tr>
<td>4.4. General conclusion</td>
<td>115</td>
</tr>
</tbody>
</table>
5. Chapter 5. Reward based-learning in subclinical checkers: The influence of feedback direction  
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1.</td>
<td>Introduction</td>
<td>116</td>
</tr>
<tr>
<td>5.2.</td>
<td>Methods</td>
<td>120</td>
</tr>
<tr>
<td>5.3.</td>
<td>Results</td>
<td>128</td>
</tr>
<tr>
<td>5.4.</td>
<td>Discussion</td>
<td>147</td>
</tr>
<tr>
<td>5.5.</td>
<td>Conclusions</td>
<td>152</td>
</tr>
</tbody>
</table>

6. Chapter 6. Reward-based learning in subclinical checkers: The influence of feedback magnitude  
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1.</td>
<td>Introduction</td>
<td>154</td>
</tr>
<tr>
<td>6.2.</td>
<td>Methods</td>
<td>159</td>
</tr>
<tr>
<td>6.3.</td>
<td>Results</td>
<td>168</td>
</tr>
<tr>
<td>6.3.1.</td>
<td>Positive study (Positive-neutral vs. positive-positive)</td>
<td>168</td>
</tr>
<tr>
<td>6.3.2.</td>
<td>Negative study (Negative-neutral vs. negative-negative)</td>
<td>184</td>
</tr>
<tr>
<td>6.4.</td>
<td>Discussion</td>
<td>201</td>
</tr>
<tr>
<td>6.5.</td>
<td>Conclusions</td>
<td>211</td>
</tr>
</tbody>
</table>

7. Chapter 7. Reward based-learning in subclinical checkers: Symptom-specific alteration of reward magnitude  
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1.</td>
<td>Introduction</td>
<td>213</td>
</tr>
<tr>
<td>7.2.</td>
<td>Methods</td>
<td>215</td>
</tr>
<tr>
<td>7.3.</td>
<td>Results</td>
<td>221</td>
</tr>
<tr>
<td>7.4.</td>
<td>Discussion</td>
<td>237</td>
</tr>
<tr>
<td>7.5.</td>
<td>Conclusions</td>
<td>244</td>
</tr>
</tbody>
</table>

8. Chapter 8. General Discussion  
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1.</td>
<td>Synthesis of the research findings</td>
<td>245</td>
</tr>
<tr>
<td>8.2.</td>
<td>Extra analysis</td>
<td>256</td>
</tr>
<tr>
<td>8.3.</td>
<td>Deficits in negative prediction errors? Discussion of general hypothesis</td>
<td>258</td>
</tr>
<tr>
<td>8.4.</td>
<td>Limitations</td>
<td>262</td>
</tr>
<tr>
<td>8.5.</td>
<td>Future directions</td>
<td>264</td>
</tr>
<tr>
<td>8.6.</td>
<td>Conclusions</td>
<td>266</td>
</tr>
</tbody>
</table>

References | 267 |
List of tables

Chapter 2
2.1. Conditional probabilities between events two previous trials ago (n-2) and the actual trial in the event sequences employed in block 2.................................................................57
2.2. Conditional probabilities between two previous trials and one actual trial for the sequences employed at block 3........................................................................58

Chapter 3
3.1. Reasons for exclusion from the laboratory based study.................................74
3.2. Checker and non-checker group similarities and differences for the OCI-R scores and DASS scores ........................................................................................................75
3.3. Exclusion criteria for each picture type based on the valence scale ratings........77

Chapter 4
4.1. Characteristics of experiments for each pilot, containing number of participants, number of trials, values of target probability between options and maximum accuracy for the two possible strategies, investigated in this study.................................................................90
4.2. DASS and OCI-R scores for participants in pilot 5...........................................91

Chapter 5
5.1. Reasons and number of participants excluded and included in the study........123
5.2. Mean, standard error and statistics for the comparison of checkers and non-checkers sample characteristics across the positive-neutral and negative-neutral experiments........125
5.3. Statistical ANOVA analyses conducted on the measures of the binary decision making task in this study.................................................................128
5.4. Pairwise comparisons for picture type in the picture validation task in the arousal scale Picture-type.............................................................................................................130
5.5. Sub-block accuracy mean values for all participants for blocks 1, 2 and 3.......135
5.6. Significant interactions (experiment vs. group) from sub-blocks two to five in block 2.................................................................................................................136
5.7. Independent post-hoc t-tests for each experiment between groups with mean accuracy levels and standard error values........................................................................137
5.8. Independent t-tests for lags 2, 4 and 5 for block 2 between groups for both experiments.................................................................................................143
Chapter 6
6.1. Reasons and number of participants excluded and included in the study………………161
6.2. Mean, standard error and statistics for checkers and non-checkers in the positive-neutral and positive-positive experiments……………………………………………………………………………………………………………………………163
6.3. Mean, standard error and statistics for checkers and non-checkers in the negative-neutral and negative-negative experiments…………………………………………………………………………………………………………………………………………………………………165
6.4. Statistical tests employed for each analysis conducted for the binary decision making task………………………………………………………………………………………………………………………………………………………………………………………………………………………………………168
6.5. Pairwise comparisons for picture type in the arousal scale………………………………170
6.6. Pairwise comparisons for the main effect of sub-block accuracy for blocks 1, 2 and 3…174
6.7. Pairwise comparisons for picture type in the arousal scale for the negative study…………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………
## List of figures

<table>
<thead>
<tr>
<th>Chapter 1.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1. Fronto-striatal OCD model proposed by Menzies et al. (2008). Simplified diagram summarizing putative regions and circuits which may be affected in OCD</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 2.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1. General study design</td>
<td>45</td>
</tr>
<tr>
<td>2.2. Examples of positive, negative, neutral and checker-related pictures used in the Picture Validation Task</td>
<td>51</td>
</tr>
<tr>
<td>2.3. Representative trials of the Picture Validation Task</td>
<td>51</td>
</tr>
<tr>
<td>2.4. Representative sketch of lab-based behavioural experiment</td>
<td>53</td>
</tr>
<tr>
<td>2.5. Adapted formula from DeGroot &amp; Schervish, (2012) to make the calculus of cross-correlation and auto-correlation</td>
<td>66</td>
</tr>
<tr>
<td>2.6. Auto-correlation for a hypothetical pattern sequence (LLRRLL…) for lags 1-5</td>
<td>67</td>
</tr>
<tr>
<td>2.7. Hypothetical values for the auto-correlation of a sequence (experimental sequence) and of the cross-correlations for two groups, calculated for lags 1-5</td>
<td>68</td>
</tr>
<tr>
<td>2.8. Hypothetical values for the auto-correlation of a sequence (experimental sequence) and the cross-correlations for two groups (maximize and pattern searcher)</td>
<td>69</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1. Mean valence ratings for positive, neutral, negative, and checking-related selected pictures displayed separately for checkers and non-checkers</td>
<td>78</td>
</tr>
<tr>
<td>3.2. Mean arousal ratings for positive, neutral, negative, and checking-related selected pictures displayed separately for checkers and non-checkers</td>
<td>79</td>
</tr>
<tr>
<td>3.3. Mean comfort ratings for positive, neutral, negative, and checking-related selected pictures displayed separately for checkers and non-checkers</td>
<td>80</td>
</tr>
<tr>
<td>3.4. Mean valence ratings for positive, neutral, negative, and checking-related considering all picture validation task studies from study 1 to 6 together</td>
<td>81</td>
</tr>
<tr>
<td>3.5. Mean arousal ratings for positive, neutral, negative, and checking-related considering all picture validation task studies from study 1 to 6 together</td>
<td>82</td>
</tr>
<tr>
<td>3.6. Mean comfort ratings for positive, neutral, negative, and checking-related considering all picture validation task studies from study 1 to 6 together</td>
<td>83</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1. Mean accuracy for the positive and negative experiments in pilot 1</td>
<td>94</td>
</tr>
<tr>
<td>4.2. Mean accuracy in each sub-block of 40 trials for the positive-positive and negative-negative experiments in pilot 1</td>
<td>96</td>
</tr>
<tr>
<td>4.3. Cross-correlation probabilities from lags 1 to 5 between the positive-positive and negative-negative experiments in pilot 1</td>
<td>98</td>
</tr>
</tbody>
</table>
4.4. Accuracy for two participants in pilot 2………………………………………………………101
4.5. Accuracy in each sub-block of 32 trials for two participants in pilot 2………………102
4.6. Cross-correlation probabilities from lags 1 to 5 for two participants in pilot 2………103
4.7. Accuracy for two participants in pilot 3……………………………………………………..105
4.8. Accuracy in each sub-block of 32 trials for two participants in pilot 3………………106
4.9. Cross-correlation probabilities for lags 1 to 5 for two participants in pilot 3…………107
4.10. Accuracy in pilot 4……………………………………………………………………………109
4.11. Mean accuracy in each sub-block of 32 trials in pilot 4………………………………110
4.12. Cross-correlation probabilities from lags 1 to 5 in pilot 4……………………………111
4.13. Accuracy in pilot 5 for each participant……………………………………………………113
4.14. Accuracy in each sub-block of 32 trials for two participants in pilot 5………………114
4.15. Cross-correlation probabilities from lags 1 to 5 for three participants in pilot 5……..115

Chapter 5
5.1. General study design for the positive-neutral and negative neutral experiments……….122
5.2. Mean comfort ratings for positive, neutral, negative, and checking-related pictures displayed separately for checkers and non-checkers………………………………………..131
5.3. Mean accuracy values for checkers and non-checkers in the negative-neutral and positive-neutral experiment for blocks 1, 2 and 3……………………………………133
5.4. Sub-block accuracy mean values for all participants for blocks 1, 2 and 3…………….134
5.5. Sub-block accuracy mean values for checkers and non-checkers, divided between experiments for block 2………………………………………………………………..138
5.6. Win-shift / Lose-shift strategies per participant divided by group and experiment……139
5.7. Comparison between the auto-correlation values (experimental sequence) and the cross-correlation values (participant) for lags 1 to 5 in block 1…………………………….140
5.8. Cross-correlation probability from lags 1 to 5 between checkers and non-checkers for the negative-neutral and positive-neutral experiment in block 2……………………141
5.9. Cross-correlation probability from lags 1 to 5 between checkers and non-checkers for the negative-neutral and positive-neutral experiment in block 3……………………….144

Chapter 6
6.1. Sub-block mean accuracy values for all participants for blocks 1, 2 and 3……………172
6.2. Win-shift / Lose-shift strategies per participant divided by group and experiment……175
6.3. Lose-shift mean values for checkers and non-checker for block 1 between the positive-neutral and positive-positive experiment………………………………………..177
6.4. Auto- and cross-correlation values for the experimental sequence and the participant’s sequence for lags 1 to 5 in block 1……………………………………………………179
6.5. Auto- and cross-correlation values for the experimental sequence and the participant’s sequence for lags 1 to 5 in block 1. Comparison between checkers and non-checkers for the positive-neutral and the positive-positive experiment..............................................180
6.6. Cross-correlation probability from lags 1 to 5 between checkers and non-checkers considering both experiments in block 2.................................................................181
6.7. Cross-correlation probability from lags 1 to 5 between checkers and non-checkers considering both experiments in block 3.........................................................182
6.8. Sub-block accuracy mean values for all participants for blocks 1, 2 and 3 for the Negative Study..............................................................................................................189
6.9. Win-shift / Lose-shift strategies divided by group and negative experiments..........192
6.10. Lose-shift mean values for checkers and non-checker for block 1 between the negative-neutral and negative-negative experiment..............................................193
6.11. Auto- and cross-correlations for lags 1 to 5 for block 1, while considering both groups and negative experiments together.............................................................195
6.12. Auto- and cross-correlations for lags 1 to 5 for block 2. Comparison between checkers and non-checkers between the negative-neutral and the negative-negative experiments.....196
6.13. Auto- and cross-correlation values for lags 1 to 5 of block 3. Comparison between the negative-neutral and the negative-negative experiments........................................198

Chapter 7
7.1. Accuracy mean values for checkers and non-checkers in the negative-neutral and symptom-related experiment for blocks 1, 2 and 3.................................................225
7.2. Sub-block accuracy mean values for all participants for blocks 1, 2 and 3..............226
7.3. Sub-block accuracy between checkers and non-checkers, considering both experiments in block 2.....................................................................................................229
7.4. Win-shift / Lose-shift strategies per participant divided by group and experiment.....230
7.5. Comparison between the correlation values of the experimental sequence and the participant’s sequence from lags 1 to 5 in block 1, considering both groups and experiments together.................................................................231
7.6. Cross-correlation probability from lags 1 to 5 between checkers and non-checkers for the negative-neutral and symptom-related experiment in block 2......................232
7.7. Cross-correlation probability from lags 1 to 5 between checkers and non-checkers for the negative-neutral and symptom-related experiment in block 3.........................234
List of Abbreviations

AC - auto-correlation
ACC - Anterior cingulate cortex
ANOVA - Analysis of variance
BG - Basal ganglia
BU - Bournemouth University
CBT - Cognitive Behavioural Therapy
CC - cross-correlation
DASS - Depression, Anxiety and Stress Scales
DASS-A - DASS anxiety scale
DASS-D - DASS depression scale
DASS-S - DASS stress scale
DBM2 - Dynamic Belief mixture model
Dif - Difference between the sequence's auto-correlation value of one lag and the cross-correlation value of this same lag
dLPFC - Dorsolateral prefrontal cortex
DM - decision making
DOSPERT - Domain Specific Risk-Taking Scale
DSM-III - Diagnostic and Statistical Manual of Mental Disorders, Third Edition
DSM-V - Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
ERN - Error-related negativity
ERP - Event-related potential
ERPt - Exposure and response prevention technique
F – female
fMRI - Functional magnetic resonance imaging
FRN - Feedback related negativity
GAD - Generalised anxiety disorder
GDT - Game of dice task
HC – Hippocampus
5-HT - 5-hydroxytryptamine (serotonin)
IA – interaction
IAPS - International Affective Picture System
IGT - Iowa gambling task
IU - Intolerance of uncertainty
IUS - Intolerance of uncertainty scale
L – Left
M – Male
ME - main effect
mdT - Mediodorsal thalamus
NPm - no money player
NPv - no videogame player
OCD - Obsessive Compulsive Disorder
OCI - Obsessive Compulsive Inventory Questionnaire
OCI-R - revised Obsessive Compulsive Inventory questionnaire
OFC - Orbitofrontal cortex
p_b - p value with Bonferroni corrections
PD - Parkinson's disease
PFC - prefrontal cortex
PIS - Participant information Sheet
Pm - money player
PTSD - post-traumatic stress disorder
Pv - videogame player
PVT - Picture validation task
R – Right
RBL - Reward-based learning
RBLT - reward-based learning task
SAM - Self-Assessment Manikin
SCR - skin conductance responses
SSRI - Selective serotonin reuptake inhibitor
UoS - University of Surrey
UPPS - Impulsivity scale
vmPFC - Ventromedial Prefrontal cortex
Y-BOCS - Yale–Brown Obsessive Compulsive Scale
# List of appendices

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix 2.1</td>
<td>Copy of ethics approval for University of Surrey studies</td>
<td>265</td>
</tr>
<tr>
<td>Appendix 2.2</td>
<td>Copy of ethics approval for Bournemouth University studies</td>
<td>266</td>
</tr>
<tr>
<td>Appendix 2.3</td>
<td>Example of online questionnaire</td>
<td>267</td>
</tr>
<tr>
<td>Appendix 2.4</td>
<td>Example of participant information sheet (PIS)</td>
<td>274</td>
</tr>
<tr>
<td>Appendix 2.5</td>
<td>Consent forms</td>
<td>276</td>
</tr>
<tr>
<td>Appendix 2.6</td>
<td>Example of Strategy questionnaire</td>
<td>278</td>
</tr>
<tr>
<td>Appendix 2.7</td>
<td>Statistical table of extra analysis, considering all experiments together</td>
<td>295</td>
</tr>
<tr>
<td></td>
<td>(analysis 1) and all negative experiments (analysis 2)</td>
<td></td>
</tr>
<tr>
<td>Appendix 2.8</td>
<td>Table resuming all significant results between groups and between experiments for accuracy in blocks 1, 2 and 3, win-shift and lose-shift and Dif of correlation values for blocks 1, 2 and 3</td>
<td>296</td>
</tr>
</tbody>
</table>
Chapter 1: Literature Review

1.1. Introduction

The aim of this thesis was to investigate potential decision making deficits in subclinical checkers. Previous research findings indicate that obsessive compulsive disorder (OCD) individuals have no deficits when faced with decision making (DM) tasks where the rules are clear and explicit to participants and where participants do not need to rely on feedback, e.g. decision making under risk tasks. However, they present deficits in reward-based learning tasks (RBLTs) involving decision making under ambiguity, where participants are blind to the rules of the task and need to use trial-by-trial feedback to learn the rules and make decisions. It is still not clear why OCD individuals have deficits in tasks that use feedback information to make predictions.

Feedback information can be positive or negative or mixed, and it can have different magnitudes. Previous studies often used mixed feedback manipulations and did only sometimes investigate feedback magnitude manipulations. This thesis would like to address this gap in the literature because it might be possible that subclinical checkers present specific deficits to only some feedback forms and / or magnitudes. In addition, it is not known how symptom-related feedback would influence OCD performance in these tasks. This was investigated in one of our studies.

Furthermore, previous studies that investigated reward-based learning (RBL) in OCD relied on the manipulation of probabilities between options, but they ignored the investigation of conditional probabilities between previous and current events. Interestingly, there is some evidence that participants usually search for conditional structures (i.e., patterns) in reward-based learning tasks (RBLTs), even when they do not exist (Unturbe & Corominas, 2007, Vulkan, 2000). This bias in behaviour might influence the performance of participants within non-pattern sequences and it could affect RBLT performance, even in tasks with random conditional probabilities. To our knowledge, no study has investigated and compared learning in RBLTs with random conditional probabilities (no pattern in the stimulus sequence but one option is more likely to win) vs. non-random conditional probabilities (pattern learning) when investigating people with OCD. Therefore, besides the need for more investigation concerning the effect of feedback magnitude on OCD performance, it is also important to consider the effect of non-random conditional probabilities in OCD performance.
Specifically, the first aim of this PhD thesis was to examine how performance is affected in subclinical checkers, while facing experiments with different positive feedback magnitudes, distinct negative feedback magnitudes or symptom-related feedback. The second aim of the thesis was to investigate subclinical checker performance in these distinct reinforcer magnitudes when facing random and non-random conditional probability environments under different feedback conditions.

The aim of the literature review presented in this chapter is to provide an overview of the background literature that is relevant for the proposed research and to discuss certain design choices for our experiment based on the background literature. First, it will focus on OCD, i.e. its symptoms and subtypes, the differential diagnosis to other related disorders (Chapter 1.2), and comorbidities with anxiety disorders and depression (Chapter 1.3). Secondly, current OCD models will be discussed (Chapter 1.4) before turning to treatment options (Chapter 1.5). Thirdly, the literature review will discuss emotional and cognitive deficits in OCD (Chapter 1.6), before focusing on decision making and reward-based learning deficits associated with OCD (Chapters 1.7 and 1.8). Here, an overview about reward based learning, associated brain structures and neurotransmitters will be provided and a review about studies involving feedback magnitude and conditional probability manipulation will be given. Based on this review, considerations for our research design planning will be discussed at the end and an overview about the research aims and hypotheses will be given (Chapter 1.9).

1.2. Obsessive Compulsive Disorder

1.2.1. Diagnosis

Obsessive compulsive disorder (OCD) is one of the most handicapping existent neuropsychiatric disorder (Menzies et al., 2008), affecting more than 50 million people in the world (Sasson et al., 1997). Ruscio, Stein, Chiu & Kessler (2010) estimated that 2.3% of the U.S. population suffers from lifetime OCD, while 1.3% of the UK population presents the condition (Torres et al., 2006).

OCD is characterized by obsessions and compulsions (American Psychiatric Association, 2013). Obsessions are persistent and repetitive thoughts that are intrusive and considered distressing and disturbing by the individual (American Psychiatric Association, 2013). They cause enhanced levels of distress and anxiety (American Psychiatric
Association, 2013). Obsessions are common in the general population and can exist in a varying degree of symptom severity, i.e. continuum from non-existent to subclinical to clinically relevant levels. Clinically diagnosed obsessions are often associated with high levels of distress with the symptoms and low quality of life due to the symptoms (American Psychiatric Association, 2013). OCD individuals usually feel driven to perform compulsions, which are repetitive behaviours or mental acts not specifically related with the nature of the obsession but they are aimed to prevent, stop or control obsessions, or reduce anxiety provoked by the obsessions (American Psychiatric Association, 2013). These compulsions are an important characteristic of OCD but they are not present in all OCD patients (Stein et al., 2010).

According to the American Psychiatric Association (2013), the nature of the obsessions and compulsions varies a lot between individuals but there are some common subtypes found in the adult population: contamination obsessions and compulsions about cleaning (washing); symmetry obsessions and compulsions about ordering or counting (ordering/counting); taboo obsessions and related compulsions (aggressive/sexual/religious obsessions) and fear of harm obsessions and security checking compulsions (checking). The most common compulsions found in the general population are checking and washing (Ball, Baer & Otto., 1996).

In the DSM-V (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition), OCD is classified within the obsessive-compulsive related disorder section. Other obsessive-compulsive related disorders are: body dysmorphic disorder, hoarding disorder, trichotillomania, excoriation disorder and obsessive-compulsive disorders induced by substance or medication (American Psychiatric Association, 2013). According to the DSM-V, intrusive thoughts (obsessions) and the use of stereotyped behaviour (compulsions) are common in all of these disorders. Many of these disorders frequently present compulsions that are body-focused (hair pulling, skin picking, etc.). For example, people with trichotillomania present the habit to constantly pick and pull hair, what characterizes a compulsion (Ferrão, Miguel & Stein, 2009). However, the presence of body-focused compulsions is not sufficient for an OCD or a hoarding disorder diagnosis (American Psychiatric Association, 2013). For example, trichotillomania is classified as a separate disease from OCD because (1) obsessions are less frequently preceding compulsions, (2) compulsions are usually preceded by sensory phenomena which are not present in OCD, and (3) compulsions are body-focused (Ferrão et al., 2009).
1. Please note, hoarding was originally considered as an OCD subtype but since the DSM-V, it is classified as a separate but comorbid disorder (American Psychiatric Association, 2013). Hoarding is usually driven by the fear of losing important objects (obsessions) and a distorted view of possessions (Saxena, 2007). Hoarding compulsions are related with rituals of saving and keeping objects that clutter the space (Saxena, 2007). There are several diagnostic reasons for diagnostically separating OCD and the hoarding disorder, e.g. earlier age of onset of symptoms, more severe family and social disability, higher indexes of anxiety and depression, and lower insight about one’s own symptoms (Saxena, 2007).

1.2.2. Common comorbidities with OCD

The most common comorbidities with OCD are anxiety (75.8%) and mood disorders (63.3%) (Ruscio et al., 2010). It is very important to consider these comorbidities when studying OCD because underperformance on neuropsychological tests could be a reflection of these comorbidities and not OCD itself (Abramovitch, Mittelman, Tankersley, Abramowitz & Schweiger., 2015). Actually, it is possible that underperformance on a neuropsychological test could be dependent on the levels of anxiety or mood disorders as depression. In fact, in this case, it would be possible to assess the influence of these comorbidity scores throughout an ANOVA that evaluates if the means of a dependent variable are equal between a group with higher scores of a comorbidity and a group with lower scores of this same comorbidity. Clearly, the study should present enough participants with both lower and higher scores, so the sample does not present excessive outliers and a similar number of participants between groups. Additionally, it would be possible to use the scores of this comorbidity as a covariate in an analysis of covariance (ANCOVA). For this, the sample should present a variation of anxiety and depression scores across participants, so participants should not be excluded for moderate levels of depression and anxiety as it has happened in this study. Indeed, these common comorbidities usually presented in OCD are discussed in this section and an investigation of the effects of these comorbidities in the obsessive compulsive sample employed in this study is discussed in Chapter 8.

1.2.2.1. Depression

Depressive disorders are characterized by the presence of sad, empty or irritable moods, persistent for more than two weeks (American Psychiatry Association, 2013).
Depression is 10 times more prevalent in OCD than in the general population (Denys, Tenney, Megen, Geus & Westenberg, 2004).

There is some support to assert that depression is likely to be a secondary symptom to OCD, meaning that depression seems to occur in the OCD population because of the distress and impairment caused by the obsessions and compulsions (Bartz & Hollander, 2006). Factors that are likely to lead to an increase of depressive symptoms in people with OCD are: severity of symptoms, number of hospitalizations, number of aggressive obsessions, number of suicide attempts, disabilities, ageing, caffeine abuse, phobias, and generalized anxiety disorder (Tükel, Polat, Özdemir, Aksüt & Türksoy, 2002). Even though depression is not considered a primary cause of OCD symptoms, it can affect the efficacy of cognitive behaviour therapy (Pallanti, Grassi, Sarrecchia, Cantisani & Pellegrini, 2011), and how patients perform in a neuropsychological task. For this reason, it is important to consider the presence of clinical depression as an exclusion criterion in research studies as this can generate incorrect conclusions about cognitive deficits in OCD.

1.2.2.2. Anxiety disorders

Anxiety is a future-oriented mood state associated with the preoccupation and planning / preparation for possible future negative outcomes, triggering physiological symptoms (stress, muscular tension, constant vigilance), cognitive mechanisms (worriedness) and certain behaviours (e.g. avoidance) to prepare for and prevent negative events (Barlow, 2004). Usually, anxiety disorders begin in childhood and are more frequent in women than men with a ratio of 2:1 (American Psychiatric Association, 2013). However, anxiety disorders differ in the nature of the situations that are feared. For instance, in social anxiety, the object of fear is social interaction, while in phobias the object of fear could be, e.g. spiders, closed rooms, or heights. Eight psychiatric conditions are classified as anxiety disorders in the DSM-V, i.e. separation anxiety disorder, selective mutism, specific phobias, social anxiety disorder, panic disorder, agoraphobia, generalized anxiety disorder, and anxiety induced by substance/medication.

Although about 75.8% of the OCD population presents a comorbidity with anxiety (Ruscio et al., 2010), OCD is not considered to be an anxiety disorder according to the new DSM-V (American Psychiatric Association, 2013). OCD is characterized by stereotyped compulsive rituals, usually associated with the obsessions and not observed in any of the anxiety disorders cited above (American Psychiatric Association, 2013). Moreover, anxiety is
It should be noted that it is not clear to what extent anxiety can be a response to the obsessions (secondary effect) or its primary cause (Stein, Craske, Friedman & Phillips, 2014). For example, anxiety might be the primary cause because OCD symptoms tend to worsen after stressful situations that generate anxiety (Stein et al., 2014). However, as anxiety is not primarily present in all OCD patients (Ruscio et al., 2010), it is possible anxiety appears in OCD as a secondary symptom because heightened levels of OCD symptoms can cause anxiety. Actually, although anxiety plays a central role in OCD (Tynes, White & Steketee, 1990), anxiety is a non-specific symptom that is common in a number of others psychiatric disorders, as depression and schizophrenia, for example (Bartz & Hollander, 2006), so it is possible anxiety is more of a comorbidity associated with OCD than its cause.

Evidence suggests that anxiety disorders such as panic disorders and generalized anxiety disorders (GAD) are more commonly associated with OCD (Pallanti et al, 2011). Firstly, the comorbidity between OCD and panic disorders ranges from 13 to 56% in the literature (Pallanti et al., 2011). Secondly, GAD’s are identified as a life comorbidity in OCD in 30% of the cases (Pallanti et al., 2011). Both disorders present overlap of symptoms as ruminative thoughts (obsessions) and distortions about probabilistic thinking and risk over-evaluation (Pallanti et al., 2011). However, individuals presenting general anxiety usually obsess about real situations, whereas people with OCD commonly obsess about alien and bizarre nature facts (Barlow, 2004).

Emerging neuroscientific evidence supports the separation of anxiety disorders and OCD. More specifically, neuroscientific models of anxiety disorders commonly focus on amygdalo-cortical (i.e. insular cortex and ventromedial prefrontal cortex (vmPFC)), and amygdalo-hippocampal interactions (Shin & Liberzon, 2010). A hyperactivation of the amygdala in the presence of feared stimuli compared to rest (Shin & Liberzon, 2010) can be found in different anxiety disorders, i.e. phobia, general anxiety, and panic disorder, as well as in the post-traumatic stress disorder (PTSD) (Shin & Liberzon, 2010). In addition, the insula cortex also shows enhanced activation in anxiety disorders which is usually related to the monitoring of internal body state (Paulus & Stein, 2006). Interestingly, the findings are different in OCD. Firstly, the findings regarding the amygdala activation in OCD are

not always present in OCD (Stein et al., 2010; Bartz & Hollander, 2006). Finally, neuroscientific evidence also supports this distinction between OCD and anxiety disorders by showing that different structures are affected in these disorders (Shin & Liberzon, 2010; Menzies et al., 2008; more details below in this section).
inconsistent (Cardoner et al., 2011), i.e. some studies have found amygdala hyperactivation (e.g. Van den Heuvel et al., 2004; Simmon, Kaufmann, Müsch, Kischkel & Kathmann 2010; Cardoner et al., 2011), others found amygdala hypoactivation (e.g. Cannistraro et al., 2004; Britton et al., 2010) or no difference in activation (e.g. Lawrence et al., 2007). Secondly, the insula cortex does not appear to mediate OCD symptoms, unlike in general anxiety disorders (Shin & Liberzon, 2010).

1.2.2.3. Post-traumatic stress disorders

The comorbidity between post-traumatic stress disorders (PTSD) and OCD is about 0.5% (Welkowitz, Struening, Pittman, Guardino & Welkowitz 2000). Since the introduction of the DSM-V, PTSD is classed as a trauma and stressor related disorder (American Psychiatric Association, 2013), and separated from anxiety disorders. PTSD can occur after exposure to traumatic events (e.g. threatened death, serious injury or sexual violence of relatives, close persons or oneself). It expresses itself in symptoms of recurrent and intrusive thoughts, distressful repetitive dreams about the traumatic event, distress when facing cues that remember the traumatic event and presence of dissociative reactions such as flashbacks (American Psychiatric Association, 2013).

OCD and PTSD are both characterized by the presence of intrusive thoughts, avoidance from objects that cause distress or the triggering of obsessions and they also perform specific behaviours to decrease anxiety (Silva & Marks, 1999). However, OCD can be separated from PTSD by the difference between their obsessions because OCD obsessions are more alien-like and not related to a real traumatic event as it is for PTSD (American Psychiatric Association, 2013).

In summary, anxiety disorders and mood disorders can be separated from OCD but they are very common comorbidities. OCD studies do not always measure anxiety or depression levels in their study samples and even fewer studies control how much differences between anxiety and depression levels between the OCD group and the control group could explain their research findings. Specially, it is important to screen the sample for these common comorbidities, excluding the possibility that these disorders would influence results and conclusions about decision making deficits in OCD. The next section will discuss cognitive and neuroscientific OCD models and related evidence.
1.2.3. OCD: One disorder or several distinct subtypes?

OCD continues to be very heterogeneous in its nature of presented obsessions and compulsions. Therefore, great efforts are made in the research community to identify meaningful subtypes that could be more homogenous, which could improve the understanding about the disease and related treatment (Saxena, 2007). Recent research showed that it seems to be the case that each subtype varies in the way they process information (Muller & Roberts, 2005). For example, Leopold & Backenstrass (2015) showed that washers performed better in distinct neuropsychological tests measuring attention, memory and processing speed than checkers. Actually, Nakao et al. (2009) showed that checkers are more impaired in non-verbal memory than washers.

Several neuroimaging studies have also shown that subtype symptoms may be mediated by partially distinct neural correlates (Van Den Heuvel et al., 2009). For example, these authors have shown that patients from the washing subtype had a smaller grey matter volume in the dorsolateral caudate nucleus, usually involved in habit learning and habit initiation than symmetry/ordering and checking subtypes. In addition, patients from the checking subtype had a smaller grey and white matter volume in the bilateral anterior temporal lobes than the other two subtypes (Van Den Heuvel et al., 2009), an area usually well connected with the orbitofrontal cortex (OFC), hippocampus formation, and the striatum (Kondo, Saleem & Price, 2005).

Based on this evidence, it might be more useful to study OCD subtypes as separate conditions. Many studies still do not differentiate between the subtypes and this could be causing inconsistencies of findings across studies, especially in the research fields of neuropsychology and clinical neuroscience (Abramovitch et al., 2015).

1.2.4. Medication and long-term treatment effects in clinical or subclinical OCD

Another limitation of many neuropsychological studies is that OCD patients are often taking medication. This medication status is usually reported in OCD literature (Abramovitch et al., 2015) but a critical reflection about the diminishing effect of some types of this medication on cognitive performance is often missing (Abramovitch et al., 2015).

A way to avoid medication effect is the recruitment of medication-naïve samples. However, these patients are more difficult to recruit and there are ethical issues related to withholding medication for research purposes. Saying that, not all OCD patients are on medication. Another alternative to avoid medication effects is the recruitment of subclinical
samples consisting of not clinically diagnosed participants with OCD symptoms. These individuals are not recruited from the clinical population and they are not seeking treatment for a specific condition (Gibbs, 1996). One can distinguish between two types of subclinical participants: individuals who present some kind of psychological difficulty but not strong enough to be considered a clinical condition; or individuals who qualify for a clinical diagnosis but are not officially diagnosed and who do not seek treatment (Gibbs et al., 1996). Usually, subclinical participants are easier to recruit compared to a clinical sample, it is more ethical (and easier to get ethical approval) to study these subclinical participants compared to naïve clinical patients where treatment is withheld for research reasons. Finally, subclinical samples do present less extreme cases of comorbidities associated with OCD (Gibbs et al., 1996), such as anxiety (Zhu et al., 2014), and therefore they give a cleaner picture about OCD itself.

In summary, OCD is a psychiatric condition, where individuals present obsessions and associated compulsions that diminish their quality of life. Usually, OCD individuals present related comorbidities, as depression and anxiety disorders, so when studying OCD one should consider comorbidities. Additionally, OCD is divided in distinct subtypes, where checking is the most common one. As there is evidence that OCD subtypes might be very different in nature and might relate to differential brain activations, it is also important to focus in only one subtype in the study of OCD. Finally, besides focusing in the exclusion of the most common comorbidities and in the study of one unique subtype, it is important to avoid possible medication effects, so in this thesis, we will investigate unmedicated subclinical checkers with mild to moderate anxiety and depression levels.

1.3. OCD Models

Different models tried to explain OCD in the literature. Previous models, such as the cognitive behavioural model by Salkovskis (1999) used psychological hypothesis to explain OCD whereas the most recent neuroscientific models tried to link OCD symptoms to brain network changes (Menzies et al., 2008; Melloni et al., 2012). These two models will be briefly explained below with an emphasis on the neurocognitive model.

1.3.1. Cognitive appraisal models

Salkovskis (1985) cognitive appraisal model was partially based on a previous model from Rachman (1981). In this model, obsessional and potential triggering stimuli prompts
intrusive thoughts that are considered alien and irrational by the individual (Sakovskis, 1985). When the individual believes these thoughts will have negative implications, obsessive thoughts arise. According to the author, when there is an element of responsibility, regarding these thoughts, they will be repeated and they become uncontrollable. This, in turn, leads to neutralizing responses, as compulsions. Actually, one of the most effective treatments for OCD, the cognitive behavioural therapy (CBT) works with the idea that encouragement about the re-interpretation of the patient’s intrusive thoughts, as presenting less negative implications, will help the patient (Salkovskis, 1999). Actually, the cognitive behavioural model was very useful in the generation of treatments for OCD, especially CBT. However, it has failed to include neurological or biological factors that could be responsible for the aetiology of OCD (Foster & Eisler, 2001). More recently, neurocognitive models were created, as a tentative to investigate the neural networks involved in this pathology, as well as how CBT and others treatments work, regarding neural networks.

1.3.2. Neuroscientific OCD models

1.3.2.1. Fronto-striatal network models

The most commonly cited neuroscientific models of OCD involve two modulated fronto-striatal circuitries. One of these models was developed by Menzies et al. (2008). Figure 1.1 below displays a modified version of their model which was mainly developed based on evidence from functional Magnetic Resonance Imaging (fMRI) and neuropsychological studies. The fronto-striatal OCD model suggested by Melloni et al. (2012) is very similar but it is more based on neuropsychological and Event-related potential (ERP) study evidence.
Figure 1.1 Fronto-striatal OCD model proposed by Menzies et al. (2008). Simplified diagram summarizing putative regions and circuits which may be affected in OCD. Yellow boxes indicate regions comprising the traditionally implicated orbitofrontal-striatal loop. Blue boxes indicate regions comprising the dorsolateral prefrontal-striatal loop. White boxes indicate additional brain regions putatively involved in OCD. Red arrows indicate connections proposed as components of frontal-striatal loops by Alexander, DeLong & Strick (1986). Green arrows indicate connections incorporated into frontal-striatal loops by Lawrence, Sahakian & Robbins (1998). Blue arrows indicate connections with supportive anatomical evidence as denoted by number (1) Cavada & Goldman-Rakie (1989); (2) Takahashi, Ohki & Kim (2007); (3) Aron, Behrens, Smith, Frank & Poldrack (2007) and (4) Mink (1996). Picture is taken from Menzies et al. (2008).

Historically, the abnormal function of the orbitofrontal-striatal circuit (Figure 1.1) in OCD was repeatedly mentioned in the literature (for a review see: Menzies et al., 2008). This circuit consists of areas such as the orbitofrontal cortex (OFC), basal ganglia (BG) nuclei, such as the ventral striatum and the ventral pallidum, and the mediodorsal thalamus (mdT). This circuit receives inputs from the anterior cingulate cortex (ACC), the hippocampus (HC), and the amygdala, the three areas ascribing affective function to this system (Menzies et al., 2008). The orbitofrontal-striatal circuit is related with motivation, monitoring and decision making and reward-based or implicit learning, habit acquisition and action thresholding (Melloni et al., 2012).
Bechara, Damasio, Tranel & Damasio (1997) demonstrated that the integrity of the OFC is not only related to motivation and emotions but is also important in feedback dependent decision making tasks with options of unknown payoff probabilities. This is supported by evidence from neuropsychological studies reporting deficits in tasks involving probabilistic feedback learning in patients with OFC lesions (Fellows, 2011). This seems to be the case, because the OFC monitors changes in feedback value (Menzies et al., 2008) that includes changes in feedback magnitude (feedback strengths), feedback probability (chance to receive a feedback) and feedback delay (immediate vs. delayed feedback) (Lak, Stauffer & Schultz, 2014).

The function of the other areas in this circuit is as follows: (1) The ACC is related with error and action monitoring, conflict detection and feedback processing (Melloni et al., 2012); (2) The basal ganglia influences the generation of motor and cognitive habits (Graybiel & Rauch, 2000); (3) The hippocampus is related with declarative and episodic memory (Phelps, 2004); (4) The amygdala is specialized in emotional conditioning and learning, esp. fear conditioning, where a neutral stimulus becomes aversive after being paired with an aversive event (Phelps, 2004). Emotional content can influence the encoding of episodic memories and the association between emotions and these memories seems to be related with the amygdalo-hippocampal interaction (Phelps, 2004).

In OCD patients, the activation of this OFC-striatal circuit is enhanced. More specifically, the OFC is hyperactive as shown in fMRI studies (Menzies et al., 2008) and in Positron Emission Tomography (PET) studies with OCD patients, either at rest or in symptom provoking paradigms (Rauch et al., 1994; Cottraux et al., 1996). In addition, OFC volume seems to be reduced in this population (Atmaca, Yildirim, Ozdemir, Tezcan, & Poyraz, 2007). Secondly, Graybiel & Rauch (2000) reported hyperactivity in the basal ganglia in OCD patients. In accordance, patients with hyperactivity in the basal ganglia or areas that project to it presented an increase in obsessive compulsive symptoms (Laplane et al., 1989). Thirdly, ACC volume is increased (Szeszko et al., 2004) and hippocampal volume is decreased in OCD (Kwon et al., 2003). Finally, some volumetric and activity differences in the amygdala were also reported but the involvement of the amygdala in OCD pathophysiology is controversial (see section 1.3.2: Anxiety disorders). More importantly, it was suggested that this hyperactivation of the OFC-striatal circuit is responsible for the decision making and RBL deficits in OCD patients, especially where emotional and reward-
based processing is important (Starcke, Tuschen-Caffier, Markowitsch, & Brand 2010; Menzies et al., 2008).

Note, however, OCD pathology is not limited by the orbitofrontal-striatal circuit. Recently, attention has been drawn to the implication of the dorsolateral PFC-striatal circuit in OCD (Figure 1.1, left panel; Menzies et al., 2008; Melloni et al., 2012). This loop comprises the dorsolateral prefrontal cortex (dlPFC), the caudate nucleus, the posterior parietal cortex and the arcuate premotor area (Alexander et al., 1986). The dlPFC and the parietal cortex have been associated with executive functions (incl. goal setting and task-switching), working memory, selective attention and response inhibition (Cabeza & Nyberg, 2000; Milad & Rauch, 2012). The dlPFC was also associated with goal setting and action monitoring during decision making (Heekeren, Marrett & Ungerleider 2008).

In OCD patients, the dlPFC-striatal circuit is hypoactive (Remijinse et al, 2006). Additionally, OCD patients have a decreased dlPFC volume (Lucey et al., 1997) and it was suggested that this hypoactivation of the dlPFC-striatal circuit could be responsible for spatial attention, working memory, and executive function deficits in this population (Melloni et al., 2012; Menzies et al., 2008; Van den Heuvel, 2005).

In summary, several fronto-striatal circuits are involved in the pathophysiology of OCD, i.e. a hyperactive OFC-striatal circuit involved in reward-based learning and decision-making and a hypoactive dlPFC-striatal circuit involved in selective attention, working memory, response inhibition and executive functions. This should be kept in mind when looking at the neurotransmitter changes in OCD and effects of pharmaceutical treatments on the activation of these fronto-striatal circuits and related OCD symptom reductions (Saxena, Brody, Schwartz & Baxter, 1998; Perani et al., 2008; Menzies et al., 2008).

1.3.2.2. Neurotransmitter changes in OCD

Many (dopaminergic, serotonergic, noradrenergic, glutamatergic, and cholinergic) neurotransmitters are involved in the activation and inhibition of the extended frontal-striatal networks (Dalley, Mar, Economidou & Robbins, 2008; Krebs, Boehler, Roberts, Song & Woldorff., 2011) and, to our ever evolving knowledge, several of them are also involved in OCD pathophysiology. (Aouizerate, Guehl & Cuny, 2005; Denys et al. 2004; Pittenger, Krystal & Coric, 2006).

The strongest evidence for neurotransmitter changes in OCD patients is linked to serotonergic and dopaminergic systems. More specifically, serotonin reuptake inhibitors
(SSRI) are given as typical treatment to OCD patients to enhance serotonin levels. In addition, dopamine D2 receptor antagonists are being used in the treatment of OCD with some success, in combination with SSRIs (Westenber, Fineberg & Denys, 2007). Actually, OCD patients present an abnormality in the dopamine-receptor binding in areas such as the basal ganglia, linked with OCD pathophysiology (Kim et al., 2007; Perani et al., 2008) which seems to reflect a dopamine hyperactivity in the striatum (Perani et al., 2008) and there is also evidence that OCD presents polymorphism (genetic variation), regarding the coding of dopamine receptor D4 (Millet et al., 2002). These evidences indicate that dopamine levels are enhanced in OCD whereas serotonin levels are reduced.

However, serotonin and dopamine are not the only neurotransmitters involved in OCD. In recent years, the interest in the role of glutamate in OCD increased (MacMaster, Oneill & Rosenberg, 2008) and according to Carlsson (2001), OCD is also a hyper-glutamatergic condition. This neurotransmitter is known to interact with dopamine and serotonin (Rosenberg, Macmillan & Moore, 2001), as glutamate receptor antagonists increase serotonin-receptor stimulation, so glutamate acts as an inhibitor of serotonin activity. However, it has the opposite effect concerning dopamine, as more dopamine also leads to more glutamate (MacMaster et al., 2008).

In fact, this complex interaction between distinct neurotransmitter systems and fronto-striatal circuits could explain OCD treatments and emotional and cognitive deficits in OCD patients. In section 1.4, a short description of OCD treatments will be given. Afterwards, emotional and cognitive deficits in OCD patients will be briefly reviewed in relation to the neuroscientific OCD model (Section 1.5.) before turning the focus to decision making and reward-based learning deficits in OCD (Sections 1.6 and 1.7).

1.4. Typical OCD treatments

Cognitive Behavioural Therapy (CBT) is the first line treatment in OCD but it is often combined with pharmacological treatments (March & Mulle, 1998; Baldwin et al., 2005). These treatment options will be discussed below.

1.4.1. Cognitive Behavioural Therapy (CBT)

CBT is the first line treatment in OCD and one of the most commonly used CBT techniques in this patient group is the exposure and response prevention technique (ERPt) (Chambles et al., 1998). In this treatment, a previous assessment about the nature of the
patient’s obsessions and compulsions, as well about her avoidance behavior, is implemented using clinical interviews or questionnaires, as the Yale-Brown Obsessive Compulsive Scale (Y-BOCS) (Skeketee, 1994). Posteriorly, the participant is exposed to the source of her anxiety, imagining it or in the presence of it, in vivo (Foster & Eisler, 2001). Afterwards, there is a response prevention component, where the patient is encouraged to do not engage in the compulsions generated by the increased anxiety (Foster & Eisler, 2001). The exposure is repeated during sessions, until the patient anxiety levels decrease (Foster & Eisler, 2001).

CBT is considered the most effective treatment for OCD with 77% to 100% effectiveness (Freeston et al., 1997; Ladouceur et al., 1995). However, CBT does present issues. There is a considerable number of patients who do not want to start therapy in the first place and the drop-out rates are high (Stanley & Turner, 1995). CBT is also less effective when OCD patients present depression as comorbidity (Cox, Swinson, Morrison & Lee, 1993). There is evidence that CBT acts on the medial prefrontal cortex (medial OFC and rostral/subgenual subregions of the ACC) (Milad et al., 2005). So, CBT seems to trigger cortical/subcortical top-down effects on OCD (Goldapple et al., 2004).

1.4.2. Pharmacological treatments

Serotonin reuptake inhibitors (SSIRs) are a well-known treatment to decrease OCD symptoms (Rauch et al., 1998). They are drugs that block serotonin (5-HT) transporters and are associated with reduction of OFC and striatum hyperactivity in unmedicated OCD patients (Saxena, 1998). However, 40% to 60% of OCD patients fail to respond to SSRI medication (Holander & Rosen, 2002) and the ones who respond, usually present reminiscent symptoms that affect their quality of life (Goodman, McDougle, Barr & Aronson, 1993), so the hypothesis that others systems than the serotonergic one are involved in the pathophysiology of OCD is very plausible (Koo et al., 2010). Indeed, antipsychotic drugs, acting on serotonergic and dopaminergic systems can benefit patients who do not respond to SSIRs only (Fineberg, Nigam & Sivakumaran, 2006; Koo, Kim & Kim, 2010). Moreover, dopamine receptor antagonists reduce symptoms of OCD, in combination with SSRIs (Koo et al., 2010). This shows that several neurotransmitters are involved in the pathology of OCD, esp. serotonin and dopamine.
1.5. Emotional and cognitive deficits in OCD

The pathophysiology of the frontostriatal circuits (Menzies et al., 2008; Melloni et al., 2012) and related serotoninergic, dopaminergic, and glutamatergic neurotransmitter imbalances in OCD patients (Krebs et al., 2011; Dalley et al., 2008) are likely to be responsible for the emotional, cognitive and decision making deficits in OCD. More specifically, the pathophysiology of the OFC-striatal circuit loop can be linked to modulations of emotional regulation and emotional perception (Lawrence et al., 1998), as well as reward-based decision making and learning that can be altered based on its feedback direction (i.e., positive vs. negative), feedback magnitude, feedback probability and feedback delay (Hikosaka & Watanabe, 2004) what could affect reward-based learning decision making in OCD. Imbalances in the dlPFC-striatal loop could be responsible for others cognitive deficits in OCD. These areas are usually related to set shifting, working memory and response inhibition (Milad & Rauch, 2012). For more details, emotional deficits, cognitive deficits and finally, reward-based decision making and learning deficits will be reviewed below.

1.5.1. Emotional deficits in OCD

Each particular emotion is associated with behavioural and physiological correlates (Panksepp, 1982). For example, finding a spider in the house whilst being afraid of spiders can result in screaming and jumping on the sofa (behavioural response) and enhance one’s heartbeat, sweating, eye pupil diameter (physiological response of the sympathetic activation system). Emotions can also impact future behaviour when learning and making decisions, and fine-tuning responses in a given emotional context (Stern, Nota, Heimberg, Holaway & Coles, 2014). Therefore, poor emotional regulation can affect our behaviour, physiology, as well as learning and decision making in many situations (Bechara, 2004).

According to Mennin, Holaway, Fresco, Moore & Heimberg (2007), three factors could lead to emotional regulation deficits: (1) a poor understanding of the meaning of a certain emotion, (2) an enhanced negative reactivity to the presence of a particular emotion (i.e., negative or positive emotions are not very well accepted), and (3) a deficit in regulating the responses related to the emotions.

People with OCD seem to have deficits related to all three factors: They have difficulty to recognize and express their own emotions, which is called alexithymia (Kang,
Namkoong, Yoo, Jhung & Kim, 2012). Secondly, Calkins, Berman & Wilhelm (2013) suggested that individuals with OCD show a hyperactive emotional response to specific thoughts (obsessions) and interpret them with a heightened negative salience and the resulting avoidance behaviour, e.g. avoiding related thoughts and situations, often leads to the occurrence of compulsions in people with OCD (Calkins et al., 2013).

Dittrich, Johansen, Fineberg & Landro (2011) suggested that all these emotional deficits could contribute to other cognitive deficits, esp. decision making. According to the authors, OCD deficits related with emotions could be responsible for the disruption of cognitive information. Interestingly, Stern et al. (2014) have shown that all OCD subtypes presented poor understanding and negative reactivity to both positive and negative emotions and they assumed this might be because people with OCD perceive emotions as a lack of control. Actually, OCD is associated with a predisposition to avoid novelty and new situations (Coles, Schofield & Pietrefesa, 2006) with a high intolerance of uncertainty (Holaway, Heimberg & Coles, 2006) and this could be linked to the avoidance of new emotions they are not prepared to cope (Stern et al., 2014). Additionally, according to Stern et al. (2014), this avoidance could lead to compulsions that would be a mechanism created to avoid or suppress thoughts that usually trigger specific emotions.

Actually, neuroscientific evidence about altered brain function related with emotional processing in people with OCD (see Chapter 1.3.2) confirms the presence of emotional deficits. As previously stated in the OCD model section (Chapter 1.3.2.1), the activation of other brain areas associated with emotional regulation is enhanced in OCD, i.e. the striatum, the ACC, and the OFC (see Menzies et al., 2008 for a review). More specifically, the OFC presented an enhanced brain activation triggered by symptom-specific stimulation (Simmon, Kaufmann, Müsch, Kischkel & Kathmann, 2010) and by general aversive stimuli ((Rolls et al., 2004; Simmon et al., 2010). In this way, these findings are in line with the assumption that the OFC is involved in the decoding of emotional values of rewards and punishments in relation to reward-based decision making and learning (Kringelbach & Rolls, 2004; Bechara, Tranel, Damasio & Damasio, 1996; Menzies et al., 2008), and that it is dysfunctional in OCD (Melloni et al., 2012).

**1.5.2. Cognitive deficits in OCD**

Cognitive impairments linked to dlPFC-striatal circuit dysfunction are set shifting, response inhibition, and working memory (for reviews see Dittrich & Johansen, 2013;
Menzies et al., 2008; Van der Wee et al., 2003 and Harkin & Kessler, 2011). In addition, memory impairments are often discussed in relation to OCD (Müller & Roberts, 2005).

More specifically, OCD participants presented deficits in a go/no-go task where participants had to inhibit responses when facing no-go cues (Bannon, Gonsalvez, Croft & Boyce, 2002). OCD participants also showed deficits in attentional and affective set shifting tasks, where participants had to inhibit response representations that were valuable in the past but no more valuable in a current state (Menzies et al., 2008). For instance, in attentional set shifting tasks, the target dimension (as colour or shape, for instance) changes across the task and participants have to adapt to these changes (Menzies et al., 2008). In affective set shifting tasks, as well as reversal learning, the feedback magnitude of a stimulus changes across the tasks (Menzies et al., 2008), so a stimulus that represented a reward before, could represent a penalty afterwards.

In addition, some forms of memory also seem to be affected in OCD patients. For example, Müller & Roberts (2005) reported that OCD patients are often unsure if they really executed a previous action and people with OCD are especially affected in their episodic memory, which is the one related with personal events, remembered from the past (Müller & Roberts, 2005). Episodic memories can be either verbal (biographical facts) or non-verbal (visual-spatial episodes) and there is great support to the hypothesis that visual-spatial memory is more affected than verbal memory in OCD patients (Müller & Roberts, 2005).

Actually, these findings could be related with others findings of hippocampus dysfunction in OCD, as the hippocampus is related with episodic memory (Phelps, 2004).

In their review, Harkin & Kessler (2011) argued that the division by modalities to classify OCD memory deficits (i.e. verbal vs. visuospatial) is not the optimal way to engage in the problem. Actually, the authors proposed that memory impairments are possibly secondary to others executive functions masked in the memory tasks employed in the investigation of memory deficits in OCD (Harkin & Kessler, 2011). For instance, if a memory task presents an attentional set shifting component, differences between groups will be found but not because of memory deficits but because of the inability to disengage from previous cognitive sets. Hence, these OCD deficits found in verbal and visuospatial memory tasks might be more related with executive deficits in OCD (Harkin & Kessler, 2011) and tasks that would increase cognitive demands would have more chance to present memory deficits (Harkin & Kessler, 2011).
In fact, the discussion above can be also transferred for RBL deficits found in OCD (see Chapter 1.6 below). Indeed, common impairments found in OCD concerning RBL tasks could also be related to components in the task that are not related with reward-based learning and decision making. For instance, working memory deficits in OCD (Müller & Roberts, 2005) could generate deficits in a RBL task where participants need to use previous results to make accurate decisions. If OCD participants are impaired in the ability to remember previous results and cues, or previous decisions, they could present a deficit to make accurate decisions, what could erroneously inform the researcher that OCD participants present a deficit in reward-based learning. In this way, the deficit would be more related with working memory than with reinforcement learning.

This could be also true for set shifting deficits associated with OCD participants (Bannon et al., 2002). Bannon et al. (2002) have shown that OCD participants are impaired in a go/no go task, as they are less able to disengage from previous cues. This deficit could influence the performance in a RBL task, if, for instance, OCD participants were unable to disengage from previous positive or negative results in the task, what could affect their learning of optimal strategies at the end.

Actually, this inability to disengage could be not only related with positive or negative cues but with symptom-related stimuli. In this way, tasks presenting symptom-related stimuli associated with the subtype of the group being studied could enhance this deficit, based on Simon et al. (2010) findings that OCD individuals are oversensitive to symptom-related stimuli. In this way, distinct reinforcers could increase or decrease deficits that could posteriorly affect the learning of optimal strategies in RBL tasks for OCD.

In fact, in the next sections, reward-based decision making and learning will be reviewed in more detail because this will form the background of the thesis. We will especially, focus on the role of feedback in these tasks.

1.6. **Reward based decision making and learning in OCD**

Bechara et al. (2005) proposed that decision making under uncertainty can be divided in decision making under risk and decision making under ambiguity. In the OCD literature, studies can be found that investigated either decision making under risk or decision making under ambiguity or both (see Starcke et al., 2010; Dittrich & Johanssen, 2013; Kim et al., 2015 and Zhang et al., 2015). These will be reviewed below.
1.6.1. Decision making under risk in OCD

Decision making under risk occurs when the probability of getting a reward or a punishment is previously known for each presented option in the task and the amount of risk of each option (for instance, the amount of reward and penalty offered for each option) is clear (Dittrich & Johansen, 2013). In this way, the probability of each outcome is known and the participant must decide between a risky choice (change of more reward but also change of more punishment) or a safer choice (Brand, Labudda & Markowitsch 2006). A safe choice presents a higher reward probability (easy to find) but a lower reward value (less money), for instance, while a risky choice has a lower reward probability (hard to find) but a higher reward value (more money, for instance) (Brand et al., 2006).

One famous task involving decision making under risk is the Game of Dice Task (GDT; Brand et al., 2006; Starcke et al., 2010; Dittrich & Johanssen, 2013; Kim et al., 2015 and Zhang et al., 2015). In this task, participants are offered a starting capital of £1000 and they are expected to maximize their winnings during the course of 18 trials. In each trial, one only virtual single dice is shown and participants should bet on the upcoming number of this dice when this dice is thrown. Participants can choose between the six possible numbers of this dice by choosing either to bet on only one number, or a combination of two, three or four numbers. After betting on one or more numbers, participants receive feedback (gain or loss of money), as the amount of money won or lost is shown in the screen. At the end of the game, participants can accumulate at most 19,000 pounds or lose at most 17,000 pounds.

Each bet, regarding each combination of numbers (for instance: only one number or two numbers and so on) is associated with a specific and previously known probability of gain or losses. These probabilities are 1:6 for one number; 2:6 for two; 3:6 for three and 4:6 for four numbers, respectively. Participants have less chance of winning when choosing just one number because the probability associated with only choosing one number is lower than the other options. Additionally, they are able to win or lose more money (£1000) if choosing only one number, compared to choices of two numbers (£500), three numbers (£200), or four numbers (£100). In this way, choosing only one or two numbers is considered as a higher risk and disadvantageous choice in the long-term because participants can win more but they also have a higher change of error, losing more. In the other hand, choosing three or four numbers is considered as a lower risk and advantageous choice in the long-term.

In addition, this task is considered as an explicit decision making task because the winning probabilities for each option (group of numbers to choose) are clear for the
participants at the beginning of the task and the amount of risk of each choice is also well known (Brand et al., 2005; Starcke et al., 2010). Indeed, as the rules of the task are explicit, participants are able to previously calculate the risk of each option (i.e., choosing from 1 number to 4 numbers) and apply strategies related to the explicit rules. However, they could also use trial by trial feedback information to help in their decisions (Brand et al., 2006).

Most importantly, people with OCD do not have any difficulty in performing this GDT where reward probabilities and magnitudes are explicit (Starcke et al., 2010). This is different for decision making tasks under ambiguity.

1.6.2. Decision making under ambiguity in OCD

In contrast to decision making tasks under risk that have explicit rules, decision making under ambiguity or RBL tasks have probabilities of a specific outcome (rewards or punishments) that are unknown to the participant. Hence, the outcome of their choices is unclear and participants have to find out how to perform these tasks by learning these implicit task rules from feedback (Brand, Grabenhorst, Starcke, Vandekerckhove & Markowitsch, 2007). Implicit learning means that it is very hard to keep track of the consequences of each choice during the task and participants cannot explicitly calculate the probabilities associated with each choice but use their feelings and hunches to direct their decisions (Brand et al., 2006; link: somatic marker hypothesis, Damasio, 1994). In this subsection we will describe and compare several decision making under ambiguity tasks, i.e. the Iowa Gambling Task (IGT), an associative feedback task devised by Starcke et al., (2010), and a reward based learning task designed by Frank, Seeberger & O'reilly in 2004. This discussion will lead to the suggested usage of a fourth decision making task under ambiguity which is called the binary choice task (Vulkan, 2000).

The Iowa Gambling Task (IGT) is the most known decision making task under ambiguity (Brand et al., 2006). It was developed by Bechara et al. (1994) and it is a card game with implicit monetary reward and punishment (Bechara et al., 1994). In this task, participants receive an initial fictional amount of gambling money, e.g. $2000, and they are asked to select a card from one of four card decks. The aim of the game is to increase the initial amount of money by choosing the advantageous card deck options and simultaneously to reduce money loss by avoiding the disadvantageous card deck options (Bechara, Damasio & Damasio, 2000). Participants are told that some decks are better than others and they have
to learn the rules implicitly throughout the use of positive (win) or negative (loss) feedbacks from trial to trial (Starcke et al., 2010).

In the task, participants do not know how many trials they will perform (usually 100 trials). The participant will receive positive feedback in 50% of all trials (winning money) and negative feedback (losing money) in the other 50% of all trials. In this way, positive and negative feedback are employed in this task and together. In addition, the feedback magnitude varies between decks. Two of the four decks (A and B) were associated with high fictional money gains after winning but even more losses. For example, participants can win $100 when choosing desk A and B in 50% of all trials, but they will also lose $250 in the other 50% of all trials. The other two decks (C and D) have lower gains but even lower losses. For example, participants can earn $50 per trial but they can lose only $25 (Bechara et al., 1994). This means, the decks with higher gains (A and B) are disadvantageous for the player in the long-term, as they will leave the participant with a negative debt. In contrast, decks C and D are more advantageous in the long-term because they will result in overall winnings. In this way, this task demands participants to employ feedback in the learning of probabilities between the four decks.

Concerning findings in OCD and other psychiatric populations, the IGT is often studied together with the GDT (Brand et al., 2006). Interestingly, several studies have shown that OCD patients do only present deficits in decision making under ambiguity such as the IGT (Starcke et al., 2010; Kim et al., 2015; Zhang et al., 2015; but see Dittrich & Johanssen, 2013 for opposite results) but not in decision making under risk tasks (Starcke et al., 2010). In the IGT, OCD participants usually choose more from the disadvantageous decks, while healthy participants learn within trials to choose the advantageous decks (Starcke et al., 2010). It is noteworthy that impaired IGT performance is usually related with OFC dysfunction (Bechara, 1996; Rolls, 2000), an area that seems to be hyperactive in OCD (Melloni et al., 2012). However, it is not entirely clear why OCD patients are impaired in implicit but not explicit decision making tasks (Menzies et al., 2008).

Indeed, note that the IGT has limitations (Kim et al., 2015; Zhang et al., 2015; Zhu et al., 2014). Different feedback factors could affect the performance in this task and it is difficult to separate them, e.g. the motivational signal of the feedback (feedback direction), feedback magnitude, feedback gain functions, immediate vs. delayed feedback (Busemeyer & Stout, 2002). For example, the preference for the advantageous decks in healthy individuals could be both because these decks produce a positive expected reward value or because it
produces more net gains over the duration of several trials (Dai, Kerestes, Upton, Busemeyer & Stout, 2015).

Starcke and colleagues (2010) suggested that people with OCD might not be able to associate the different feedback types (positive or negative feedback values) with the feedback probability of each option in the IGT; potentially because of their emotional deficits. It is possible that OCD is unable to access the gain-loss frequency information throughout the task because of an emotional misperception, regarding trial by trial feedback. Based on this assumption, Starcke et al. (2010) developed a simple associative feedback task to investigate this further in OCD patients. More specifically, participants had to choose between two symbols on a computer screen (triangle or circle). Afterwards, feedback appeared on the screen and participants were asked to use the feedback to learn about target probabilities and therefore to improve their performance. In all three blocks of this task, one of the options was always better than the other one (Starcke et al., 2010) but the blocks did differ in terms of the feedback direction. More specifically, in one block participants could only win points and one option always offered more points (positive feedback condition), in another block participants could only lose points and one option always offered more losses (negative feedback condition) and (3) participants could win and lose points in each option but one option had higher gains (50 to 80 points) and lower losses (-10 to -40 points) and the other option had lower gains (10 to 40 points) and higher losses (-50 to -80 points) (mixed condition). The positive and negative feedback block had 32 trials each whereas the mixed block had 34 trial presentations. Starcke et al. (2010) compared the performances of OCD patients and controls in this associative feedback task with performances in the classic IGT and GDT (Starcke et al., 2010). The main differences between the associative feedback task and the IGT were (1) the associate feedback task did not present a conflict between immediate and long-term rewards, and (2) it evaluated how participants used positive, negative, or mixed feedback to learn the optimal strategies in the task.

Findings showed that OCD patients presented deficits in the IGT but not in the GDT (Starcke, Tuschen-Caffier, Markowitsch & Brand, 2009; Starcke et al., 2010) and they also presented a lower frequency of choices for the advantageous option (deficient learning) in the associative feedback task, independent of the three different feedback conditions (Starcke et al., 2010). This could suggest that the feedback type (positive, negative, mixed) did not differentially influence patient performance when explaining OCD deficits in these decision making tasks. Thus, it seems OCD is not able to use feedback direction information or to
differentiate emotional signals to make assumptions about future events in an uncertain environment. However, while this conclusion is tempting, Starcke et al. (2010) task only had 32-34 trials per block which might be not enough to learn the task. Indeed, OCD participants could be slow learners compared to controls and it is possible they would reach the learning plateau in the learning curve later.

A third probabilistic reward-based learning task was developed by Frank (2004) for behavioural experiments with Parkinson’s disease patients. Afterwards, this task was adapted for ERP experiments (Frank, Woroch & Curran, 2005). For simplification, this task will be called the Frank task in the thesis. In the Frank task, training blocks with mixed positive and negative feedback were followed by test blocks without feedback. In each trial, participants were shown one of three pairs of Japanese letters (AB; CD; EF). These letter pairs were presented in random order. The task of the participant is to receive as much positive feedback as possible by learning to choose the better option for each letter pair. Each of these letters had a distinguished winning probability. In the AB pair, A targets had a 80% chance of winning and B target had a 20% chance of winning. In the CD pair, the chances were 70% vs. 30%, and they were 60% and 40% in the EF pair. After each letter choice, participants received correct or incorrect feedback (Frank et al., 2005). In this task, healthy participants learned that A, C and E are more likely to be linked to positive feedback than B, C and F. Most importantly, the test blocks were designed in such a way that it was possible to assess whether participants learned more by remembering items in the pairs that were more linked to positive feedback (A, C, E) or by remembering items that were more linked to negative feedback, which is called a response bias.

In the famous Parkinson’s disease (PD) study, Frank (2004) found that the response bias was modulated by dopamine medication in these patients. PD is a neurodegenerative disorder primarily associated with motor symptoms such as bradykinesia / akinesia, tremor, and rigidity (Ryterska, Jahanshahi & Osman, 2013) and caused by the degeneration of dopamine-producing neurons in the substantia nigra pars compacta (Ryterska et al. 2013) which modulates the activity of basal ganglia structures and, hence, several cortico-striatal circuits (Ryterska et al., 2013). The sensorimotor circuit is primarily affected in PD which explains the motor symptoms but there are also changes in the dlPFC-striatal circuit and the orbitofronto-striatal circuit function linked to more cognitive and decision making dysfunction (Ryterska et al. 2013). Interestingly, decision making in RBLTs is especially affected in PD when they are required to learn from positive and negative feedback.
(Knowlton, Mangels & Squire, 1996; Shohamy et al., 2004; Frank et al., 2004) but not by tasks where feedback was not necessary for learning. This is interesting, as OCD seems to show the same pattern. Frank et al. (2004) compared the performance of on- and off-medication PD patients and they found that PD patients off medication (less dopamine) learned more from negative feedback, while PD patients on medication (more dopamine) learned more from positive feedback. They suggested that this happens because dopamine overdose from medication increased the amount of dopamine bursts, encouraging learning from positive feedback, while dopamine depletion in OFF medication PD patients, increased the amount of dopamine dips encouraging learning from negative feedback.

Gründler, Cavanagh, Figueroa, Frank & Allen (2009) and Cavanagh, Gründler, Frank, & Allen (2010) investigated the same samples of OCD patients with low and high level symptoms with a similar task adapted from Frank et al. (2005) but, rather remarkably, they did not find any group differences in response times and accuracies in the learning phase, the reaction times and response biases in the test phase, and no correlations between OCD symptoms and task behaviour. However, they reported a reduced ERN in the OCD participants with higher symptom levels. In this way, OCD deficits concerning RBL are contradictory.

In summary, the previous subsections showed that OCD patients can perform decision making under risk tasks but they show deficits in reward-based decision making and learning tasks. However, the findings are partly contradictory, e.g. learning behaviour is diminished in OCD in the IGT (Starcke et al., 2010; Kim et al., 2015; Zhang et al., 2015) and in an associative feedback task (Starke et al., 2010) but not in the Frank task (Gründler et al., 2009; Cavanagh et al., 2010)). In this way, it is also still unclear why OCD presents deficits in RBLTs, i.e. which feedback related factors are relevant (feedback direction, magnitude, gain function, delay). These gaps in the literature will be addressed in the next subsection.

1.6.3. Gaps in the literature and new directions

Firstly, many decision making tasks under ambiguity (reward-based learning tasks) investigated many task-related factors at the same time and, hence, it is unclear which factors can explain the OCD deficits and which factors can be excluded. A prime example for this is the IGT which is most commonly used in the literature.

Secondly, Starcke et al. (2010) made an initial attempt to separate effects of distinct feedback directions (positive, negative vs mixed feedback) on RBL and found a diminished
performance in OCD patients independent of the feedback direction (positive negative, or mixed feedback block). This might be the case because (1) they did not have sufficient trial numbers per block, (2) they had carryover effects between blocks in their within participant design, or (3) feedback magnitude manipulations were not strong enough. It should be noted that the Frank task used by Gründler et al. (2009) and Cavanagh et al. (2010) did use mixed feedback in the learning phase and the preferred usage of positive and negative feedback is only inferred from the behaviour in the test phase of the task which does not allow interpretations about reward-based learning in a purely positive or negative feedback task (versus absence of feedback).

Thirdly, no investigation about OCD sensitivity to symptom-related feedback was employed and this is important to address, as Simon et al. (2010) have shown that symptom-related stimuli, as well as general aversive stimuli, trigger increased fronto-striatal activation in OCD and this could reflect an emotional hyperarousal to these stimuli, In this way, if general aversive stimuli, as negative feedback could influence OCD decision making, it is also possible that symptom-related stimuli will also affect it in a distinct or similar way.

Finally, one problem that is present in all the RBLTs described above is that these tasks do not address the issue regarding conditional probabilities. Conditional probabilities are the probabilities of an actual outcome in dependence of the occurrence of previous outcomes in the stimulus sequence (Aslin, Saffran & Newport 1998). Whilst people seem to employ conditional thinking, there are no reward-based decision making and learning tasks that addressed this issue in OCD. The problem is that these tasks assume that participants are only using trial by trial information during their decisions while a lot of studies have shown that participants usually make inferences about the dependence of current and previous outcomes (Vulkan, 2000 for a review). Even more interesting is that studies point out to the fact that the search of dependencies between trials influence the performance, even if the sequence is totally random (Vulkan et al., 2000; Unturbe & Corominas, 2007; Victorino, Feher da Silva, Seiss & Baldo, unpublished). In fact, if participants usually use dependencies between trials to make decisions (whether imaginary or not), possible deficits to assimilate conditional probabilities could contribute to the generation of decision making deficits in tasks with random sequences (no dependencies) such as the IGT and the Frank task.

One potential direction for future research is decomposing feedback direction and magnitude information in RBL tasks when investigating people with OCD, in addition to the investigation of the influence of symptom-related stimuli in OCD reward-based learning. A
second potential direction is to investigate how OCD deals with conditional probabilities. One way to do this is to design a task where the conditional probabilities governing the dependency between current and previous trials are non-random. In this way, it is possible to investigate feedback direction and magnitude effects in a task where either target options (e.g. option A: 80% correct; option B: 20% correct) and / or conditional probabilities are present (random vs. non-random sequences). Please note, deficits in the usage of conditional probabilities could influence RBL performance in OCD, even if a task does not present dependencies between options because people are biased to search for patterns in the environment (Jones & Pashler, 2007). In this way, it is important to address both issues: feedback magnitude manipulation and the investigation of the influence of conditional probabilities in OCD decision making. In the next subsection, the binary decision making task will be introduced. Because of its simplicity (only 2 options to choose from) this task allowed us to study the effect of feedback direction and magnitude manipulations and conditional probabilities on RBL in subclinical checkers and non-checkers. In this way, physiological evidence of feedback direction and magnitude, as well as sequential effects (conditional probabilities) will be discussed in more detail after the binary decision making task.

1.6.4. A different reward based learning (decision making under ambiguity) task: The Binary Decision Making task

A fourth typical RBLT is called binary decision making task (Vulkan, 2000). This task will be used in the present studies of this thesis. Note, this task is very similar to the associative feedback task used by Starcke et al. (2010). However, this task has not been used with OCD samples yet. In this task, participants have to decide between two options (e.g. two cards) in each trial based on their intuition. Their task is to learn via trial-and-error which option is more advantageous. Usually, each option was linked to a specific reward probability (e.g. option A: 80% correct; option B: 20% correct). The position of the target will depend on a random trial sequence operating with fixed probabilities, which are independent of the history of the outcomes and the behaviour of the participant. Generally, people are good in learning the probabilities in these tasks but they usually match responses with the probabilities associated with the task (probability matching), so if the task presents an option with 70% correct feedback, participants will choose this option with a 0.7 frequency and the other with 0.3 (Vulkan, 2000).
Actually, the advantage of this simple RBLT is that many factors can be manipulated independently, e.g. feedback direction of motivational signals or magnitude, feedback delay, or target vs. conditional probabilities. In addition, other analysis, apart the common use of accuracy, can be more easily applied to evaluate task performance. For example, in all RBLTs positive feedback is a sign of continuation (Van Duijvenvoorde, Zanolie, Rombouts, Raijmakers & Crone, 2008). After receiving positive feedback, participants are more likely to keep using the previous option (win/stay strategy). However, negative feedbacks are a sign of discontinuation from the previous planning strategy (Van Duijvenvoorde et al., 2008), and participants might try the alternative options (lose/shift strategy). In this way, participants learn to adjust their behaviour based on previous positive and negative feedback (Zanolie, Van Leijenhorst, Rombouts & Crone, 2008). This is called the win stay / lose shift strategy (Paulus et al., 2001; Evenden & Robins, 1983) which indicates how strong participants are affected by the immediate probabilities linking the actual trial to the previous outcome. Moreover, there is evidence that participants condition their decisions not only on the outcomes of the last trial, but also to other previous trials (Nicks, 1959; Anderson & Whalen, 1960; Peterson & Ulehla, 1965; Tang, 1996; Vulkan, 2000 for review). This phenomenon is called sequential effect (Wilder, Jones & Moser, 2009) and will be more discussed in Chapter 1.8. Actually, win/stay/lose/shift and sequential effects can be analyzed in these binary decision making tasks too and this is important, as currently, studies of reward-based learning in OCD only consider the use of accuracy as a measure of learning the optimal strategy in a task.

In this way, to improve knowledge for the investigation of feedback direction and magnitude, as well sequential effects, theoretical models of reward-based learning in relation to feedback direction; feedback magnitude manipulations and sequential effects (conditional probabilities) will be discussed in more detail in the next subsections.

1.7. Feedback direction and magnitude manipulations in RBL

In behavioural learning psychology and in biopsychology, reward (appetitive system) and punishment (aversive system) have been distinguished as two discrete dimensions (Fiorillo, 2013). However, physiological and computational models have treated them as opposite sides of a single continuous dimension of feedback value / magnitude (Fiorillo, 2013).
Specifically, in the single-dimension hypothesis dopamine neurons represent both, positive and negative feedback directions and are the detectors of the “goodness” of one specific event. Thus, they would emit an enhanced signal (phasic dopamine bursts) if an event was better than expected which encourages continuation; no signal, if an event is the same as expected (no phasic dopamine output changes) or a reduced signal (phasic dopamine dips) if an event is worse than expected which encourages a discontinuation / change of plan (Matsumoto & Hikosaka, 2009; Frank et al., 2004). In this way, in this model, feedback direction and feedback magnitude are on the same continuum.

The findings by Frank et al. (2004) do support the one dimension hypothesis. They have shown that OFF medication Parkinson disease patients (less dopamine) and ON medication Parkinson disease patients (more dopamine) present distinct ways of learning from positive or from negative feedback. They suggested that this happens because dopamine medication increases the likelihood of phasic dopamine bursts, encouraging learning from positive feedback. In contrast, dopamine depletion in OFF medication PD patients, increased the amount of phasic dopamine dips. Hence this study supports the idea that dopamine is both responsible by positive and negative prediction errors (see Chapter 1.6.2 for more details).

Note, generally, decision making in RBLTs is affected in PD when they are required to learn from positive and negative feedback (Knowlton et al., 1996; Shohamy et al., 2004; Frank et al., 2004).

Based on this one dimension hypothesis, one could expect that OCD patients with a hyperactive dopamine production (Perani et al., 2008) would learn more in experiments within a positive feedback direction compared to experiments within a negative feedback direction. However, as described before, Gründler et al. (2009) found no difference in accuracy, reaction times and reward bias between high subclinical OC participants and low subclinical OC participants which is a puzzling finding. Therefore, it is still unclear if OCD will present an advantage towards positive feedback.

There is also the possibility that positive and negative feedbacks are modulated by two distinct systems, i.e. the appetitive and the aversive system. According to this two dimension hypothesis, dopamine neurons are only responsible for reward signals, while others neurotransmitters might be responsible for punishment signals (Fiorillo, 2013). Evidence for this hypothesis comes from an animal study with single cell recordings, by Fiorillo (2013), showing that dopamine neurons seem to be insensitive to aversive stimuli and the variation of aversive stimuli magnitude but not to appetitive stimuli. They found initial
evidence for the presence of four different kinds of neuron groups representing distinct kinds of feedback values: reward ON; reward OFF (absence of reward); aversive ON and aversive OFF (absence of punishment). Dopamine is the neurotransmitter for reward ON neurons but others kinds of neurons as serotoninergic, noradrenergic, and cholinergic ones could be matches for the others neuron groups, even though evidence for this is still unclear (Fiorillo, 2013).

However, evidence for the two dimension hypothesis comes from fMRI studies that show activations in distinct brain areas for positive and negative feedback. (Van Duijvenvoorde et al, 2008). For example, Van Duijvenvoorde et al. (2008) showed that representation of positive and negative feedback changes within development. More specifically, they compared the responses of healthy adults (25 years) and children (8-9 years) to positive and negative feedback in a functional MRI study. Behaviourally, they found that adults and children performed more accurately after positive than after negative feedback. In addition, 8-9 years old children had more difficult to learn from negative feedback than from positive feedback compared to the adults. Additionally, the neuroimaging results revealed two distinct systems computing positive (mainly the caudate and interconnections) and negative feedback (mainly the dIPFC in adults) but it is still not clear if dopamine neurons are responsible for the representation of both feedback directions.

In summary, positive and negative feedback seem to be computed, either by one system or by two distinct systems (one vs. two dimension hypotheses). Actually, evidence indicates that OCD presents a hyperactivation of the dopamine system, so based on the one dimension hypothesis, this would decrease the probability OCD participants present dopamine dips after aversive stimuli, generating deficits in OCD concerning aversive stimuli. Therefore, in theory, based on the one dimension hypothesis, it is possible that OCD patients will have a specific deficit in the use of negative feedback but no deficits in the use of more positive feedback. In contrast, it would be expected that OCD patients might learn better after positive feedback only because their dopamine system is hyperactive (Perani et al., 2008).

However, it is well documented that dopamine modulates positive feedback but it is still not clear if dopamine also modulates negative feedback (Fiorillo, 2013). Additionally, considering that positive and negative feedback seem to activate distinct regions in the brain (Van Duijvenvoorde et al, 2008), it is possible that the two dimension hypothesis is more probable than the one dimension hypothesis. In this way, the present study hypothesis will be based on the two dimension hypothesis.
Indeed, considering the hyperactivation of the dopamine system in OCD and the two dimension hypothesis, it is possible that OCD would present a better or distinct performance after positive feedback, as it was considered for the one dimension hypothesis. In the other hand, if OCD presents deficits towards aversive stimuli, as it was already shown by Simon et al. (2010) study, it is possible that other neurotransmitter systems are affected in OCD. Actually, this would not be a surprise, as the serotoninergic (Westenber et al., 2007) and the glutamatergic systems (MacMaster et al., 2008) are affected in OCD. Indeed, the serotoninergic system was pointed out as a possible candidate for the generation of negative prediction errors (Daw et al., 2002), so if OCD is affected towards aversive stimuli within a RBL task, it is possible that negative prediction errors are affected due to deficits in the serotoninergic system of OCD..

1.8 First order and second order (sequential) effects in RBL

Wilder et al. (2009) stated that sequential effects can be divided into two categories: first and second order effects. In first order effects, response only depends on the probability of the actual stimulus (base-rate probability) and its comparison with the immediate previous stimuli (prediction errors – see subsection 1.8.1). In this way, through first order effect, individuals use feedback information to estimate option values. In second or higher-order effects, individuals learn about the correlation structure between actual trial and previous trials (i.e., relationships between one current trials and two, three…n previous trials). In this case, not only the feedback of the current stimulus is relevant but also information regarding patterns generated in previous trials (Wilder et al., 2009). Second and higher-order order sequential effects often occur without awareness of the participant and even occur when participants consciously know there are no dependencies/patterns between actual and previous trials (Vulkan, 2000). In this way, it seems that sequential effects are an important and robust cognitive bias in individuals (Wilder et al., 2009).

Usually, RBLT studies do not investigate second or higher-order sequential effects as a variable that possibly could affect decision making behaviour. They usually use random sequences of events. However, cognitive biases caused by first and second order effects and their relation to in RBL were addressed in Wilder et al. (2009)’s Dynamic Belief mixture model (DBM2). According to the DBM2, the learner assumes that the stimulus at a given trial is dependent on the first order effect probability (feedback probability) and on the second order effect probability (conditional probabilities between options). Both factors are
estimated together but while first order probabilities are rated based on feedback information, second order probabilities are rated based on previous stimulus presentations. They are not related to feedback but to the stimulus pattern in a sequence of events (Wilder et al., 2009).

For example, within the binary decision making task, participants make predictions based on the direction and magnitude of the feedback offered by one option. If the immediate previous feedback is positive, participants do not correct their strategy for the next trial (continuation) but if it is negative, participants correct their strategy and could shift options (discontinuation). In this case, this information is related with first order effects. However, based on the DBM2 model, participants also estimate dependencies between trials to make predictions about the presence of a pattern. If they believe a pattern exist, then their next decision will not necessarily depend on the feedback from their immediate previous decision but instead on the pattern regularities they predict for this event sequence. Clearly, if the decisions will depend on the pattern imagined by the participant, participants will sometimes ignore immediate previous positive and negative feedback information, while relying on their model about the pattern they imagined for this sequence. In this way, first and second order effects could also be linked to the famous exploration / exploitation dilemma (Sutton & Barto, 1999), where participants are faced with the dilemma of either ignoring immediate previous feedback information and continuing to exploit previous well known strategies (potentially suboptimal) or to trust the recent feedback information and explore new risky (potentially more profitable) strategies (Cohen, McClure & Angela, 2007; see also Chapter 1.6.2 for more details).

According to Sutton & Barto (1998) exploitation is defined as the participant’s ability to prefer and repeat previous actions that returned a good amount of reward in the past. However, to discover such actions, participants should explore new and unknown actions that were not previously selected. In this case, the participant’s ability to select unknown actions as a way to discover new and more profitable moves is called exploration (Sutton & Barto, 1998).

The dilemma and the use of exploration and exploitation strategies changes with learning (Cohen et al., 2007). Exploration behaviour can help to improve the estimates for each RBLT option over time (Sutton & Barto, 1998). This is especially important at the beginning of the task where participants are still trying to find out reward values for each option (Sutton & Barto, 1998). For instance, in the IGT, initial exploration is more effective than exploitation. If the participant explores different strategies, ignoring the initially enticing
immediate bigger rewards, the participant is more likely to be successful in the long-term (Bechara et al., 2000).

However, shifting/exploration between choice options is not always the favourable strategy. For example, in a binary decision making task without sequential effects where one option presents more targets than the other one (Vulkan, 2000), the best and optimal strategy would be to repeatedly choose the option with the highest probability (Vulkan, 2000). In the binary decision making task literature, this strategy is called maximization, as the perseverance of choices in the option with the highest probability would maximize the rewards received by the participant (Vulkan, 2000). Indeed, maximization is a particular case of exploitation, where participants repeatedly choose the option with the highest reward probability in a binary decision making task. However, it is possible, for example, that one group or participant chooses to exploit a particular strategy that is not related with the repetition of decisions in only one option, as maximization. For instance, if a task presents sequential effects (patterns), participants would be able to exploit and repeat this pattern after learning it via exploration. In this case, the exploitation of the pattern would not be called maximization as the repetition of choices are not happening in only one option but the exploitation of the pattern would still mean that participants are trying to maximize their profits by repeating theirs choices in such pattern structure.

Actually, in a binary task, maximization is the optimal strategy, as exploitation of a pattern would be in a task with sequential effects. However, it is very common that young adults do not employ this strategy and prefer to consistently shift between options, exploring new solutions (Vulkan, 2000). Indeed, one possible explanation of why participants use sub-optimal exploration in this specific task is that, as RBLTs are ambiguous in the sense that participants never know about the actual values of each option, i.e. participants cannot be sure if an option is the one with the greatest winning value in the long-run, most participants would balance between the exploration of new options and exploitation of the previously discovered simpler strategy (Sutton & Barto, 1998).

Interestingly, participants usually know that one option presents more targets (Vulkan, 2000), if questioned about this after the task. However, they would feel the urge to not ignore other possible strategies that could be better in the long-term. According to one of the main hypothesis trying to explain this phenomena (see Vulkan, 2000 for a review), this would be related with the participant’s belief about the possibility of the presence of patterns in this sequence, so participants would continue to explore in the search of new evidence of the
presence of patterns in the task (Vulkan, 2000). In this way, this would indicate that participants really estimate first and second order effects for each option, even if there are no real dependencies between actual and previous trials. A lot of studies have shown that participants usually make inferences about the dependence of current and previous trials (second order effects), even when these correlations do not exist (Nicks, 1959; Anderson & Whalen, 1960; Peterson et al., 1965; Tang, 1996; Vulkan, 2000 for review).

If the relationship of first and second order effects is intrinsic to the decision making and RBL process, irrespective of trial dependencies in the task, then it is possible that second order deficits contribute to the generation of decision making and RBL deficits in tasks without need to rely on dependencies. Hence, if one group with sequential effect deficits is compared with other group without these deficits, decision making and RBL differences could be found, even if conditional probabilities are totally random in the task. Therefore, one problematic issue of RBL studies in OCD is that they rely only on the investigation of first order effects. However, it is also important to investigate how OCD deals with second order effects.

In summary, RBLTs investigate how people deal with the emotional and rational processing of reward and punishment (wins and losses; Dittrich & Johanseon, 2013). As participants are not sure about the estimated reward values of the options, they either can explore all options (esp. at the beginning of the task) or exploit the reward value of the currently chosen option. In addition, it is possible that second order effects also influence participant’s decisions. A more detailed discussion of first and second order effects as well as related experimental data and physiological correlates will follow.

1.8.1. Physiological correlates of first order effects: Prediction error monitoring and related behaviour modulations

Prediction errors are conceptualized as the comparison between expectation and reality (Maia, 2009) and they are used to update the participant’s internal model about the optimal strategy on a trial-by-trial basis (Maia, 2009). Therefore, they can be considered as a good model for first order effects. They are affected by the feedback direction (positive or negative) as well as the feedback magnitude (e.g. Maia, 2009).

Positive prediction errors occur when the outcome is better than expected by the participant’s idea of optimal strategy and negative prediction errors are present when the outcome is worse than expected (Maia & Frank, 2011). In the case of positive prediction
errors, both the estimated value of the option and the preference to keep choosing this same option increase (continuation), while negative prediction errors have the opposite effect, increasing the odds of shifting to other unexplored option (Maia, 2009). In this last case, this kind of avoidance, increased by the probability of shifting is known as no-go learning (Frank et al., 2004). In addition, when the expected value is the same as the actual outcome, there is no prediction error at all, as no differences were found between expected and actual outcomes (Maia, 2009). In this case, it is very likely that participants keep choosing the previous action and do not explore alternative options. In other words, the strategy is not corrected because the predicted outcome was the actual outcome. However, when participants experience a negative prediction error, esp. an error, they might explore other options but note that when a previous error is represented by the absence of a reward, there is no negative prediction error (Schultz, 2010).

One model involving reward-based learning and prediction errors is the actor-critic model (Barto, 1995; Barto, Sutton, & Anderson, 1983). To understand this model, let’s consider a non-random binary RBLT, where there is at least one option whose estimated payoff value is greatest at any time (probabilities are different than 0.5). According to this model, participants can select options based on their estimated payoff values (Maia, 2009). These estimates help the participant to create a map about the possible values of each future action and to develop a strategy about the best actions to perform in each given future trial (Sutton & Barto, 1998). The actor-critic model tries to sketch the optimal way a hypothetical participant would work to find the optimal decision-making strategy and to estimate the value-function of the current strategy that is being used before an actual decision is done. It consists of two components: the actor and the critic, where the critic, stores and learns the current value of an action via prediction error calculation and the actor compares prediction error values of each past action-outcome, learns the best associations and implements the strategy to be employed in the future by the decision maker (Maia, 2009).

According to Daw, Niv & Dayan (2005), the actor role is believed to be related to structures as the dorsal striatum, whereas the critic role is related to the activation of dopaminergic OFC-striatal and meso-limbic circuits, as the ventral striatum, the amygdala and the OFC (Maia, 2009; Geisler et al., 2007).

The dopamine innervated OFC-striatal circuit (with strong connections to the ACC and the amygdala; Schoenbaum & Roesch, 2005; see Menzies et al. 2008 for more details) seems to be critical for the calculation and interpretation of prediction errors (McClure, Berns
Montague, 2003; Bray & O’Doherty, 2007; Niv & Schoenbaum, 2008). These areas seem to work together to implement the critic as they were shown to present activity during the expectation of reward (see Schoenbaum, Chiba & Gallagher, 1998 for OFC; Belova, Paton & Salzman, 2008 for amygdala).

Additionally, dopamine signals seem to reflect prediction errors by signing what is a good decision to continue or what is a bad decision to discard (Schultz, 1998, Montague, Dayan & Sejnowski, 1996; Montague, Hyman & Cohen, 2004; Schultz & Dickinson, 2000). Dopamine bursts are related to reward prediction errors, driving reinforcement learning (Fiorillo, 2013) whereas dopamine dips might be linked to negative prediction errors. Note, however, more recent research seems to indicate that other neurotransmitters might be engaged in the signaling of negative prediction errors (Fiorillo, 2013), for example, serotonin level changes (Daw, Kakade & Dayan, 2002). Both neurotransmitter functions are modulated in OCD.

Evidence from electroencephalogram (EEG) studies strongly suggests that prediction error monitoring is altered in OCD patients (Zhu et al., 2014). The event-related potential (ERP) often used to study feedback learning and, hence, learning from prediction errors is the feedback-related negativity (FRN; Holroyd & Coles, 2002). The FRN is a negative deflection that appears at fronto-central recording sites, peaking approximately 200-300ms, following feedback presentation and it is larger after the presentation of negative feedback (Holroyd & Coles, 2002). The FRN presents a distinct sensibility to the direction of the feedback presented (positive vs. negative) and to the degree of feedback uncertainty (O’Toole, Weinborn & Fox, 2011). Thus, the FRN is typically enhanced for negative feedback compared to positive feedback (Nieuwenhuis, Holroyd, Mol & Coles, 2004); it is also enhanced in the presence of unexpected outcomes compared to expected outcomes (Holroyd, Krigolson, Baker, Lee & Gibson, 2007).

The FRN is generated in the anterior cingulate cortex (ACC) (Holroyd & Coles, 2002) and It is assumed that it is a reflection of the ACC receiving reinforcement learning signals from the midbrain dopamine system to update the decision making models generated by individuals during RBLT’s (Holroyd & Coles, 2002). Indeed, Holroyd & Coles (2002) asserts that the anterior cingulate cortex (ACC) function is related with the adaptation of behavior by comparing actions with their consequences. In this way, it is possible that activity in the ACC reflected by the FRNs is also affecting others areas connected to the ACC and related with reinforcement learning signals, e.g. the OFC (Menzies et al., 2008).
As stated previously, neuroimaging studies have shown that OCD individuals have a higher activation of the OFC-striatal loop and the ACC (e.g. Menzies et al., 2008). Researchers hypothesized that this hyperactivity is generating a malfunctioning in the performance monitoring system of OCD (Pitman, 1987) because changes of activities in these areas are usually correlated with performance and error monitoring (Balleine, Delgado & Hikosaka, 2007).

Interestingly, the FRN amplitude is diminished for OCD patients compared to controls in RBLTs after receiving negative feedback from an incorrect response (Endrass, Koehne, Riesel & Kathmann, 2013; Grundler et al., 2009). Endrass et al. (2013) suggested that an attenuated feedback monitoring could be responsible for the reduced FRN amplitude in OCD patients. In addition, O’Toole et al (2012) measured FRNs for both correct and incorrect feedback trials in a modified deterministic four-choice reversal learning task and they noticed that amplitudes for FRN were similar for correct and incorrect trials in the OCD group with higher symptom scores, while FRN amplitudes were higher for incorrect choices and lower for correct choices in the OCD group with lower symptom score group; showing again that the FRN, and hence feedback monitoring of negative feedback, decreases with enhanced OCD symptomatology.

In summary, FRNs are reduced in OCD patients and this suggests a reduced monitoring and processing of prediction errors related to negative feedback. Therefore, the specific deficit in decision making under ambiguity in OCD patients is, at least partly, related to prediction error monitoring and processing dysfunction. Until now, no study has investigated how different feedback directions and magnitudes affect prediction error monitoring in OCD. This is also beyond the scope of this thesis.

1.8.2. Physiological correlates of second order effects: Conditional probability processing

Jones & Pashler (2007) argued that predictions based on past experiences drive our choices in everyday life. Therefore, it is reasonable to assume that the brain is highly adaptive when generating long-term predictions, rather than only associating experiences that occur in close temporal proximity (as prediction errors). However, even though computational theories have accepted the existence of longer-term temporal associations (Sutton, 1988; Montague et al., 2004), there is an absence of experimental efforts to investigate the influence of dependencies across trials on decision making (Jones & Pashler, 2007). Nevertheless,
several studies have demonstrated that healthy participants can detect non-random conditional probabilities and use this information to make assumptions about next decisions (Fiser & Aslin., 2001; Hunt & Aslin, 2001; Jones & Pashler, 2007). They are influenced by conditional thinking, even when these dependencies do not really exist in the task (e.g. in random sequences).

For instance, the underperformance in Jones & Pashler (2007) non-random binary decision making tasks could be explained with pattern search behaviour, even if there is no pattern to be found in the sequence (see Vulkan, 2000 for a review). For example, Victorino et al. (unpublished) examined in a recent unpublished study the effect of ageing on RBL accuracy and strategies while using two simple binary decision making tasks with positive feedback. In one task, probabilities for each target option differed and options were presented in random order. Hence, the optimal strategy was maximization. In the other task, target probabilities between options were equal (50:50) and maximization was no potential strategy. However, conditional probabilities generated a pattern in the sequence. Findings showed that while adults did not employ maximization in the first random sequence task, this same group performed better in the second pattern sequence task and reached maximum accuracy. However, 70-75 years old elderly presented a deficit in the pattern recognition but showed better maximization in the random sequence than young adults. Interestingly, this data supports the idea that maybe there is a relationship between conditional thinking and performance in a random task, which is supported by Huettel, Mack & McCarthy (2002) findings where participants admitted to see and follow pattern structures during a task, even though researchers told them that the task was totally random before the experiment. This is not surprising because most sequences in everyday life contain a mix of regular patterns and unpatterned random sequences.

There is evidence, for example, that the prefrontal cortex (PFC), the basal ganglia (specifically the striatum) and the ACC are activated during pattern recognition (Huettel et al., 2002; Zweig et al., 2002). For example, the interruption of pattern structures is correlated with the activation of the PFC and the ACC (Huettel et al., 2002). According to Bechara et al., (2005), the striatum and the ACC would be responsible for the pattern recognition of simple pattern structures, as repetition or alternation. However, more complex patterns would need others structures to help to make more accurate predictions, e.g. the OFC according to the somatic marker hypothesis (Damasio, 1994). All these circuits are affected in OCD (e.g. Menzies et al., 2008). Hence, it is possible that OCD presents deficits in the recognition of
complex patterns, especially if these patterns are constantly interrupted in uncertain (noisy) environments.

Thus, to address these issues, this present thesis had the aim to investigate about pattern recognition deficits, where three sequences, random or with a pattern, were investigated in five distinct experiments. Indeed, to investigate the effect of feedback direction, magnitude and symptom-related stimuli in OCD reward-based learning five distinct experiments were designed to represent distinct emotional stimuli.

Specially, as it is unclear if OCD presents a specific deficit regarding feedback direction and/or magnitude (either associated with symptom-related feedback, positive or negative feedback), an investigation about these issues will be addressed, while considering sequences with patterns and no-patterns, as it is possible that these deficits will distinctively affect learning in one of these sequences.

1.9. Summary and overview of experimental chapters

People with OCD have many cognitive and emotional deficits which are related to activation changes in the OFC-striatal, dPFC-striatal and mesolimbic circuits. This thesis will concentrate on implicit reward-based decision making and learning deficits in RBLTs (also called decision making under ambiguity) in subclinical checkers and non-checkers to bring more understanding to OCD decision making and learning deficits. Examples of these RBL tasks (decision making tasks under ambiguity) are the IGT, the associative feedback task, the Frank task, and the binary decision making task. Some of these tasks investigate several feedback related factors at the same time, e.g. feedback direction, feedback magnitude, feedback gain, feedback delay. This makes interpretation about the specific deficits rather difficult. Therefore, it is important to investigate these different factors in RBL whilst measuring feedback factors in a systematic way.

The aim of the studies presented in this thesis will be to compare unmedicated subclinical checkers and non-checkers (1) to investigate the effects of feedback direction (positive vs. negative) and manipulation of feedback magnitude within positive and negative feedback direction, (2) to study differential effects of disease specific feedback magnitude alterations with symptom-related pictures (checking pictures) and (3) to investigate the effect of systematic changes in target probabilities vs. conditional probabilities on learning in the RBLTs (random event sequence vs. pattern learning). For these studies a binary decision
making task will be employed because it allowed us to manipulate all these factors independently.

Concerning the feedback direction manipulation, the reward-based decision making and learning literature does not usually discriminate between the effect of positive and negative feedback types on learning and often mixed feedback is given to the participants during learning. However, it is possible that positive and negative feedback are modulated by either the same or different brain or neurotransmitter systems in the brain or by extreme variations of dopamine neuron activation (one vs. two-dimension hypotheses). Therefore, the investigation of subclinical checkers in learning environments with either positive vs. absent monetary feedback or negative vs. absent monetary feedback could bring new insight about RBL processes in OCD. This question will be addressed in Chapter 5. This first investigation, represented by studies 2 and 3, is similar to the Starke et al. (2010) study but they used two positive or two negative feedback options with no absent feedback and fewer trials. Starke et al. (2010) found OCD patients presented deficits in performance compared to controls for random target sequences.

In Chapter 6, the feedback magnitude in the positive and negative feedback experiments was enhanced to investigate the effect of feedback magnitude enhancements in positive and negative reward settings. This was done by showing emotional vs. neutral pictures shortly after the monetary feedback was given. Positive emotional pictures were presented in the positive-positive monetary feedback experiment (study 4) after winning (compared to neutral pictures), and negative emotional pictures were used in the negative-negative monetary feedback experiment (study 5) after losing (compared to neutral pictures). The alternative option was always absent monetary feedback and neutral pictures.

In Chapter 7, symptom-related emotional picture feedback was given in the negative monetary feedback experiment (study 6) in order to investigate deficit-related to feedback magnitude alterations inducted by symptom-related pictures. These pictures should have more negative emotional valence (and comfort) ratings in checkers compared to non-checkers.

In all experiments presented in Chapters 5-7, random and pattern event sequences (conditional probability manipulation) were presented in three different RBL blocks with a certain level of uncertainty (ambiguity). The first block presented random sequences with differential target probabilities (Option a: 72%, Option B: 28%) but no second order conditional probabilities. Hence, first order effects were analyzed here, i.e. win shift /lose
shift probabilities. In the second block, conditional probabilities were not random (pattern sequence) and target probabilities between the options were at chance level (50:50). Here, second order effects on the RBL of subclinical checkers was investigated. Finally, in the third block, non-random conditional and target probabilities were used to investigate which strategy (pattern search vs. option maximization) would be preferred by subclinical checkers and non-checkers, especially if there are deficits regarding first or second order effects.

This blockwise manipulation of conditional probabilities was integrated in the study design because this allowed us to study differential effects of first order and second order probabilities in subclinical checkers and non-checkers. As discussed above, RBLT studies usually employ accuracy to investigate deficits in OCD. However, it is possible that distinct combinations of responses are hidden by same amounts of accuracy. In this case, as this present work is interested in possible deficits, regarding predictions errors (first order effect) and pattern recognition (second order effect), it is important to go beyond accuracy and investigate how participants use previous feedback to make decisions. To investigate first order effects and exploration/exploitation bias, win/stay/lose/shift analysis will be employed, where the probability a participant shift after making a mistake or a correct choice will be calculated. Besides, to investigate pattern recognition, the probabilities participants are correlating one actual response to previous outcomes will be calculate (cross-correlation analysis).

As participants are never sure about the probabilities of the task, they face the dilemma between exploring new strategies (especially in the presence of patterns) or to exploit well known strategies that are safer but sometimes possibly less beneficial. Healthy young adults usually balance exploration with exploitation. However, this balance might be affected in subclinical checkers, either because participants are unable to use current feedback information to adapt their exploration / exploitation behavior (first order) or because they are unable to find possible patterns in a probabilistic event sequence (second order). These could be some of the underlying causes for the probabilistic reward-based decision making and learning deficits previously observed in OCD. As prediction errors are possibly affected in OCD, It is expected that these deficits will be more present in negative conditions than in positive conditions, as OCD seem to be hyperactive in the production of dopamine.

In this case, if the hyperactivation of dopamine in OCD decreases the probability of dips of dopamine, which in turn are responsible for negative prediction errors, OCD individuals will present a deficit in using error information to make predictions about future
decisions but no deficits in using information about rewards. This, in fact, will be more evident in negative conditions, where the magnitude of the error is higher than in a positive condition, where the errors are represented by the absence of feedback.

In truth, if dopamine dips are absent at all in OCD, this group would not present any kind of sensibility to errors irrespective of the error magnitude of the task, because there would not be any negative prediction errors at all. In this case, OCD would present deficits in distinct feedback magnitudes, as errors would never be used as information in RBLTs. However, it is also possible that OCD still presents some kind of response towards errors, and this would depend on the error magnitude of the task. Actually, OCD would be able to generate negative prediction errors but not after the error magnitude surpasses a specific threshold. In this case, it is possible that OCD will present more deficits in experiments with higher error magnitudes and less deficits in conditions with lower error magnitudes.

If OCD patients do have deficits with negative feedback, one question that remains is about how the apprehension of patterns in blocks 2 and 3 (second order probabilities) will be affected, specifically in the negative monetary feedback experiment. More specifically, if pattern learning depends more on inferences from the event sequence and less on immediate feedback information, subclinical checkers might not be affected by negative feedback in these blocks or less affected by the feedback magnitude manipulation. Potentially, checkers might rely more on the pattern information to make decisions in these pattern sequence blocks. They could learn and memorize the pattern structure, based on feedbacks at the beginning, and they could be less influenced by disruptions within the sequence and immediate feedback information. In this way, subclinical checkers would present a bias towards exploitation of the pattern structure in conditions with negative feedback and this should result in greater accuracy levels in subclinical checkers compared to non-checkers in Block 2, and potentially also in Block 3. This would not be the case for block 1 (sequence without patterns), as there are no patterns to focus on. In this way, performance in blocks 2 and 3 would not be affected or would be higher for checkers (exploitation of the pattern structures) but differences between groups would be more evident in block 1, in experiments with higher error magnitudes.
Based on the literature review, the main predictions for the thesis studies are:

1. If checkers do not present any kind of sensibility to errors, checkers will present a distinct performance than non-checkers in all the negative experimental conditions, irrespectively of the error magnitudes.

2. Checkers will not present deficits in the experiments within the positive direction, if the negative prediction error hypothesis is correct. However, they could present deficits in these experiments if (1) they present a deficit to use previous positive feedback to make decisions (2) they present a deficit in the use of previous errors associated with the absence of reward information. In this case, this deficit will possibly not be related to negative prediction errors, based on Schultz (2010).

3. If checkers do still present some sensibility to error information related to negative feedback, checkers will present less deficits in experiments with lower feedback magnitudes than higher feedback magnitudes, so performance will be similar between checkers and non-checkers in the first experiments but possibly distinct in second experiments.

4. Feedback magnitude will affect performance in pattern and non-pattern sequences, depending on how errors affect checkers. When errors do not affect checkers, they will not be affected in sequences with sequential effects, so no group differences will be found. When errors affect checkers, however, they will not rely so much on immediate and previous feedback information as non-checkers, so while non-checkers will be affected by the eventual disruptions in the pattern and shift between options more often (exploration), checkers will rely on others strategies, as the repetition of the previously learned pattern structure presented in the sequence (exploitation of the pattern). In block 1, it is possible that checkers will be more inclined to repeat responses at the same option (maximization), also ignoring target changings between options. In this way, checkers will present an exploitation of the pattern and of the option with the highest feedback probability, what will increase accuracy and correlations between previous trials for checkers. Checkers will outperform non-checkers.

5. Symptom-related emotional picture feedback will negatively affect the performance of subclinical checkers in the binary decision making task because checkers are as affected by this stimuli than by general aversive stimuli. Clearly, this effect will only be present, if checkers do present a deficit when exposed to negative stimuli.
Chapter 2: General Methods and Data Analysis

2.1. Introduction

This thesis presents a set of studies with a very similar methodology. This special methods chapter was written to describe the methods employed in most studies to avoid repetition across chapters and to be a guide. All studies contain experiments with usually two parts: an online screening questionnaire and a laboratory-based part which includes a picture validation task, a binary choice decision making task disguised as a gambling task, and several laboratory-based questionnaires. This chapter will entail a detailed description of these experimental parts. In addition, it will contain the proposed data analysis used for the studies. Particularities regarding each study and experiment will be described in the respective chapters.

The content of the respective experimental chapters is as follows: Chapter 3 describes the findings from a picture validation task (study 1) that was conducted to assess the emotional properties of 125 positive, negative, neutral or symptom-specific pictures while comparing checkers and non-checkers ratings for these pictures. Based on these values, 20 positive, 20 negative, 40 neutral, and 20 symptom-related pictures were selected for the main experiments. Chapter 4 describes the pilot experiments that tested and adjust parameters used in the newly designed binary decision making task to find a more optimal version for the main experiments. Chapter 5 describes two experiments where differential effects of feedback direction were investigated in checkers and non-checkers. The first experiment (study 2) used positive vs. absent monetary feedback and neutral emotional pictures (positive-neutral experiment), the second experiment (study 3) used negative vs. absent monetary feedback and neutral emotional pictures (negative-neutral experiment). In Chapter 6, the question of the effect of feedback magnitude on the checker and non-checker behaviour was investigated. The feedback magnitude was enhanced by introducing emotional pictures in the binary decision making task in two further experiments, i.e. the positive-positive experiment (study 4) and the negative-negative experiment (study 5). Emotional pictures were always presented in the present feedback conditions. The performance in these experiments was then compared to the performance in the experiments already reported in Chapter 5. In Chapter 7, the performance of checkers and non-checkers in a binary decision making task with symptom-related pictures and negative monetary feedback (negative-symptom-related experiment – study 6) was compared to the performance in a binary
decision making task with negative monetary feedback and neutral pictures (negative-neutral experiment) in order to investigate disease-related feedback magnitude alterations in the context of symptom-related pictures.

2.2. Experimental and overall procedures

All experiments received favourable ethical opinion from the University of Surrey Ethics Committee (Appendix 2.1 for Chapters 3, 4, 6, 7) and Bournemouth University Ethical Committee (Appendix 2.2 for Chapter 5). Figure 2.1 shows a general outline of the experiments. Each experiment was composed of two parts. The first part was an online screening questionnaire that screened for potential checker and non-checker participants for the laboratory-based study. Selected participants were invited to the laboratory part of the study which generally consisted of a picture validation task, a binary decision making task, and some laboratory based questionnaires. An exception was the study reported in Chapter 3, where participants only took part in the picture validation task in the laboratory.

![General study design](image)

*Figure 2.1 General study design. All studies consisted of two parts: an online screening part and a laboratory-based part. The specifics of the laboratory based tasks for each study will be described in the relevant chapters. A few pilot studies did not follow this exact format.*

2.3. Participant Payment

For filling out the online screening questionnaire, participants were put into an Amazon voucher prize draw (1 x £30, 2 x £10 Amazon voucher). Alternatively, participants could receive lab tokens towards the Psychology degree for completing the online screening questionnaire. The number of lab tokens varied across studies because the value for one token
per hours spent in one general study varied between universities. At University of Surrey, 0.5 lab tokens were offered at the online part and at Bournemouth University 0.25 credits were granted.

As a compensation for their time when taking part in the laboratory-based part of the experiment, participants received either a payment or lab tokens. The amount of payment varied across studies. At University of Surrey, participants received either £5 (Picture validation task – Chapter 3) or £10 (pilot studies in Chapter 4; experiments in Chapters 6 and 7). At Bournemouth University, participants received £12 (Chapter 5). The amount of lab tokens was 2 lab tokens for the experiments reported in Chapter 5. In addition, participants could win money in the binary decision making task (see Section 2.5.3.4 for details)

2.4. Part 1 - Online Screening Questionnaire

2.4.1. General description

The online screening questionnaire was used to screen for subclinical checkers and non-checkers who could take part in the laboratory-based study. This online screening questionnaire was developed using the online platform Qualtrics (see online questionnaire example in Appendix 2.3). It was used in all experiments described in Chapters 5-7. Basically, online questionnaires were the same across studies. The only difference regarded specific information about the study shown at its first two pages and information at the Consent Form (varied between Universities).

Potential participants could access the online questionnaire using a web link that was published on posters, e-mail lists, and on the SONA research platform. Basically, the questionnaire contained an information sheet, a consent form, a descriptive questionnaire asking participants about gender, age, nationality (Chapters 5-7), videogame playing habits, playing for money habits, and previous and current diagnosed mental illnesses such as depression, anxiety, panic disorder and PTSD, the revised Obsessive Compulsive Inventory questionnaire (OCI-R; Foa et al., 2002), and the Depression, Anxiety and Stress Scales (DASS; Lovibond & Lovibond, 1995).
2.4.2. Detailed description of the online questionnaires

Demographic questionnaire

Questions about age, gender, videogame playing habits, playing for money habits, and previous and current diagnosed mental illnesses such as depression, anxiety disorders and PTSD were included in this questionnaire.

Participants were asked about their videogame and their gambling for money experience because videogame practice was shown to positively affect performance in other attention demanding and perceptual learning behavioural studies (Green, Li & Bavelier, 2010 for a review; Mishra, Zinni, Bavelier & Hillyard, 2011 for attention demanding tasks). In addition, the habit to play for money was considered as a potential confounding variable as our task demanded gambling money. These variables were used to investigate possible links between gaming and gambling experience and decision making performance if they were different between checker and non-checker groups.

Obsessive-Compulsive Inventory – Revised (OCI-R; Foa et al., 2002)

The OCI-R is a comprehensive self-report measure for assessing symptoms of OCD (Foa et al., 2002) and is a simplified version of the OCI (Foa, Kozak, Salkovskis, Coles & Amir, 1998). This questionnaire was chosen because participants could easily understand and self-report symptoms online, without the need of the experimenter’s presence. It was shown that the OCI-R is an efficient measurement to discriminate OCD behaviour and different OCD sub-types (Foa et al., 2002). It contains 18 questions rated on 5-point Likert scales, evaluating the degree of distress in relation to typical OCD symptoms (Foa et al., 2002). Each of the items is rated from “not at all” (coded as 0) to “extremely” (coded as 4). Each of the six subtype subscales constituted of three questions (Foa et al., 2002). These subscales were: checking; washing; hoarding; ordering; counting and obsessing. Foa et al. (2002) reported good to excellent internal consistency for each subscale. Excellent internal consistency was found for the total score on the OCI-R (control group, $\alpha = .89$; OCD group $\alpha = .81$). Good construct validity has been reported for the OCI-R and all its subscales (Abramowitz & Deacon, 2006; Foa et al., 2002).
Depression, Anxiety and Stress Scales (DASS)

The Depression, Anxiety and Stress Scales (DASS; Lovibond & Lovibond, 1995) is an easy self-report measurement of depression, anxiety and stress. For this reason, it was chosen to evaluate these symptoms in the online screening questionnaire. This questionnaire consisted of 42 questions rated on 4-point Likert scales from “did not apply to me at all” (coded as 0) to “applied to me very much or for a good part of time” (coded as 3). The DASS has three subscales: DASS Depression (DASS-D), with scores ranging from normal (0-9 points) to very severe (28+); DASS anxiety (DASS-A) with scores ranging from normal (0-7 points) to very severe (20+); and DASS Stress (DASS-S) with scores ranging from normal (0-14) to very severe (34+). Cronbach’s alphas for the DASS-D, DASS-A, and DASS-S subscales were $\alpha = .97$, $\alpha = .92$, and $\alpha = .95$, respectively; showing excellent internal consistency. Construct validity of the DASS is also supported by studies with non-clinical and clinical samples (Antony, Bieling, Cox, Enns & Swinson, 1998).

2.4.3. Inclusion and exclusion criteria for the laboratory-based study

Participants were primarily selected for the checker and non-checker groups using the OCI-R checking subscale score. Participants were considered as checkers when they had an OCI-R checker subscale score of $\geq 5$, based on previous literature (Foa, et al., 2002; Hajcak, Huppert, Simons, & Foa, 2004) and as non-checkers when they had a score of $\leq 2$. Other exclusion criteria were (1) age (inclusion: 18-32 years) because Victorino et al. (unpublished) has shown that aging can affect decision making strategies in binary decision making tasks, (2) a current diagnosis of depression, anxiety disorders, or PTSD for checkers and non-checkers because these are common comorbidities with OCD (Ruscio et al., 2010). Finally, participants were excluded on the basis of their DASS subscale scores, i.e. DASS depression score $> 13$ (mild depression), DASS anxiety score $> 14$ (moderate anxiety), or DASS stress score $> 25$ (moderate stress). All other variables from the online screening questionnaire were measured to describe the sample but not as inclusion or exclusion criteria, e.g. gender, gaming and gambling experience.

2.4.4. Data Analysis of the Online Questionnaire

After the participants have been selected based on the online questionnaire and inclusion / exclusion criteria described above, statistical analysis was conducted on the data from participants that did take part in the laboratory study for each group and experiment.
The following dependent variables were considered: age and gender (categorical); OCI-R total, OCI-R checking, OCI-R washing, OCI-R hoarding, OCI-R ordering, OCI-R counting, and OCI-R obsessing; DASS depression, DASS anxiety, and DASS stress, playing videogames (categorical), playing for money (categorical).

Firstly, skewness and kurtosis z-scores were calculated for each variable within each group (checkers and non-checkers) and experiment to test the assumptions of normality. These assumptions were violated when z-scores were > ±1.96 (Field, 2009). In that case, non-parametric test equivalents were chosen as statistical tests wherever possible. Note, we continued to use ANOVAs where necessary even if some variables were non-parametric because of the absence of an equivalent non-parametric analysis that could be applied without much loss of data information. Khan & Rayner (2003) showed that ANOVA is robust enough to be employed in non-normal distributions and that it is more appropriate than non-parametric tests with small samples (n < 30).

Secondly, for the effects of age, of each DASS sub-scale and each OCI-R sub-scale two-way between-participant ANOVAs with the factors Group (checkers and non-checkers) and Experiment were conducted to explore group sample differences between experiments. Levene tests were performed to test for the homogeneity of variance assumption. Independent sample t-tests were performed as post-hoc tests and Bonferroni corrections were applied where necessary. Effect sizes are reported for all thesis. The effect size calculation was based on Field (2009). For categorical variables, Pearson Chi-square tests were employed or, in the case that more than 30% of the cells had expected count of less than 5, a 2-sided Fisher’s exact tests were calculated. There was no interaction analysis for categorical data.

2.5. Part 2 – Laboratory-based Study

2.5.1. Overall Procedure

After being selected for the laboratory-based part of the study based on their online screening questionnaire responses, participants were invited to come to the Psychology department of the university. All pilot studies reported in Chapters 3 and 4, as well as experiments reported in Chapters 6 and 7 were recorded at the University of Surrey whereas experiments described in Chapter 5 were recorded at Bournemouth University.

In all studies participants received the Participant Information Sheet (PIS) about the study and signed the study-related consent form. For the experiments reported in Chapters 5-
7, the laboratory-based part of the study consisted of three parts: the picture validation task, a decision-making binary choice task disguised as a card gambling task, and finally, the completion of five laboratory-based questionnaires. The study in Chapter 3, only used the picture validation task and the pilot experiments described in Chapter 4 only used the binary decision making task.

At University of Surrey, experiments were presented in an experimental research booth on a ViewSonic 1080p full HD monitor (Monitor height: 14 cm; Brightness: 5.0; Contrast: 6.0, as displayed on the monitor). At Bournemouth University, experiments were presented in an experimental research booth on a HP Elite display E231 monitor (Monitor height: 14 cm; Brightness: 90; Contrast: 80 as displayed on the monitor). Tasks were programmed in E-Prime Version 2.0 software (Psychology Software Tools, Pittsburgh, PA). At both universities, participants responded with key presses on a HP keyboard.

2.5.2. Picture Validation Task

The picture validation task was an adapted version of the Self-Assessment Manikin (SAM) questionnaire used by Lang, Bradley & Cuthbert (2008) which has been previously used to validate pictures in the International Affective Picture System (IAPS) database (Lang et al., 2008). The questionnaire was adapted in two ways. Firstly, 100 pictures were selected for the studies (details below), i.e. 50 neutral (mean IAPS valence: 5.13 ± 0.19), 25 positive (mean IAPS valence: 7.6 ± 0.32) and 25 negative (mean IAPS valence: 3.58 ± 0.16) pictures. In addition, 25 checking-related pictures were included. They were chosen from the Simon et al. (2010) picture data set. In their study, Simon et al. (2010) showed positive correlations between specific checking-relevant OCD pictures and OCD fear responses but they did not mention any possible psychological harm or discomfort of OCD participants participating in their study (Figure 2.2). Secondly, participants rated the valence (smiling/happy vs. frowning/unhappy), arousal (calm / relaxed vs. excited / aroused) and comfort (comfortable vs. uncomfortable) levels for each picture. These ratings were measured on 9-point Likert scales. Valence and arousal scales were previously used in the IAPS validation (Lang et al., 2008). However, the comfort rating was newly added in this study because studies have shown that compulsions are used by OCD patients to decrease discomfort caused by obsessions (Veale et al., 1993; Veale, 2007 for a review). In this way, comfort ratings could be an indirect measure of the presence of obsessions caused by emotional stimuli in the task.
Figure 2.2 Examples of positive, negative, neutral and checker-related pictures used in the Picture Validation Task

The picture validation task had a total of 375 trials (3 x 125 pictures). These were presented in two blocks of 186 trials (block 1) and 189 trials (block 2). At the beginning of the task, participants had two practice trials to familiarise themselves with the task.

All pictures were randomly presented on the computer screen with a ghost-white background (248, 248, 254) until participants responded by pressing the keyboard numbers 1-9 to rate the pictures (for examples see Figure 2.3). At the University of Surrey, the picture size was 8.5 cm x 14 cm and the scale size was 2.8 cm x 27.3 cm (Chapters 3, 4, and 6). At Bournemouth University, the picture size was 9.2 cm x 15.2 cm and the scale size 3 cm x 29.2 cm (Chapter 5).

Figure 2.3 Representative trials of the Picture Validation Task. From left to right: (A) valence scale trial, (B) arousal scale trial, and (C) comfort scale trial.

2.5.3. Binary decision making task

2.5.3.1. General Procedure

All experiments used a binary choice decision making task disguised as a gambling task. In this task, participants were presented with two card options over the course of many
trials. Their task was to choose one card (initially based on gut feeling) and they received feedback afterwards. Their task was to learn from trial-by-trial feedback to improve their winnings or stop their losses depending on the task type.

For the thesis, five experiments (studies 2-6) with distinct feedback profiles were designed. These feedback profiles did not change throughout the experiment. More specifically, learning was always reinforced by monetary feedback (money winnings or money losses) and / or by emotional pictures (positive, negative, symptom-related). In two experiments (positive-neutral experiment (study 2) and positive-positiv e experiment (study 4)), positive monetary feedback was given when choosing the correct card (winning 2 pence), whereas no money was given when choosing the incorrect card (absence of feedback). In three experiments (negative-neutral experiment (study 3), negative–negative experiment (study 5), and negative-symptom related experiment (study 6)), negative monetary feedback was given when choosing the incorrect card (losing 2 pence), whereas no money was won when choosing the correct card (absence of penalty).

The second form of feedback was the presentation of a picture which was by default neutral, except in the positive-positive experiment, where a positive picture was presented when the correct card was chosen; in the negative-negative experiment, where a negative picture was displayed when choosing the incorrect card and in the negative-symptom related experiment, where a symptom-related picture was presented after choosing the incorrect card (more details in Chapters 5 (study 2 as positive-neutral and study 3 as negative-neutral); 6 (study 4 as positive-positive and study 5 as negative-negative) and 7 (study 6 as symptom-related experiment)).

All experiments were programmed using E-Prime v2.0. Participants sat in a comfortable chair in front of the computer screen in a silent room with medium luminosity. In addition to written instructions (Appendix 2.4 for an example of Participant information sheet and 2.5 for Consent form), participants learned the task by seeing a demonstration of each display screen within a trial which consisted of 5 display screens (Figure 2.4). Each screen was explained by the experimenter. In this way, participants knew what to expect in each trial. Furthermore, two practice trials were given. Finally, the binary decision making task was presented. It consisted of three blocks with 160 trials (total trials 480).
2.5.3.2. Stimuli and Trial Structure

In a typical binary decision making task trial, a black fixation cross (University of Surrey (UoS): 0.5 cm x 0.5 cm; Bournemouth University (BU): 0.7 cm x 0.7 cm) was displayed on the centre of the screen on the computer monitor with a medium green background (RGB 34, 139, 34) for the entire trial. In addition, the reverse sides of two identical virtual playing cards (UoS: 3.5 cm x 5.5 cm; BU: 4 cm x 6 cm) were presented left and right from the fixation cross (fixation cross-card distance: 6 cm; card-screen edge distance: 14 cm) until the participant chose one of these cards (Figure 2.4a). Participants chose the left card by pressing the keyboard key “a” and the right card by pressing the keyboard key “l”. To avoid confusion, sticker labels with the letters “L” (L = left side card) and “R” (R = right side key) were attached to these response keys.

After responding, the outline of the chosen card turned blue (RGB 0, 0, 255) for 1000 ms (Figure 2.4b). This way, participants could remember the chosen card without using too much working memory, which reduced working memory deficits as a potential confounding variable (see Harkin & Kessler, 2011 for a review on working memory deficits in OCD).

![Figure 2.4](image)

Figure 2.4 Representative sketch of lab-based behavioural experiment. From left to right: A) the participant chooses a card; B) the outline of the chosen card turns blue; C) feedback is disclosed; D) a picture is displayed; E) Finally, the fixation point reappears.

Afterwards, both cards turned around showing the target (money bag) and the non-target (upside-down money bag) for 1000 ms (Figure 2.4c). In the positive monetary feedback experiments, participants received 2 pence when choosing the correct card with the target, but nothing when choosing the incorrect card with the non-target. In the negative
monetary experiments, participants received no reward when choosing the correct card with
the target, but lost 2 pence when choosing the incorrect card with the non-target. Participants
were instructed about this monetary reward assignment at the beginning of the task.

After the cards disappeared, a randomly assigned picture (UoS: 10.5 cm x 5.5 cm;
BU: 11 cm x 6 cm) was displayed for 1000 ms (figure 2.4d). The emotional valence of this
picture varied between experiments. Generally, the displayed picture was neutral. Exceptions
were that (1) a positive picture was displayed after a correct card choice in the positive-
positive experiment, (2) a negative picture was presented after an incorrect card choice in the
negative-negative experiment, and (3) a symptom-related picture was shown after an
incorrect card choice in the symptom-related experiment. All the pictures used at the binary
decision making task were previously selected on the basis of the picture validation task
study reported in Chapter 3. Pictures were always randomly chosen from picture sets of 2 x
20 neutral pictures (Mean valence: 4.48 ± 0.15), 20 positive pictures (Mean valence: 2.21 ±
0.14), 20 negative pictures (Mean valence: 7.25 ± 0.15), and 20 symptom-related (Mean
valence: 5.68 ± 0.17) pictures.

At the end of the trial, a fixation cross was presented for 1000 ms (figure 2.4e). Note,
information about the amount of money won / kept so far was given every 32 trials.

2.5.3.3. Blockwise manipulation of event sequences using target option and
conditional probabilities

This blockwise manipulation of event sequences was the same for all five experiments
presented in Chapters 5-7. These event sequences were designed to create distinct
probabilistic learning environments that varied between the three distinct learning blocks.
Each block contained a predetermined event sequence of the target positions which varied
based on the target option probability and the conditional probability (random vs. pattern).
The exact manipulation of these probabilities will be described below.

2.5.3.3.1. Designing the block event sequences based on the Markov chain logic

The block sequences were developed and manipulated using the probabilistic model
of Markov chains. According to Markov (1971), Markov chains are postulated when the
emergence of one specific event (for example, target emergence in the left side) is dependent
on the emergence of other previous events (for example, previous target on the right side).
The degree of dependence can be manipulated by increasing or decreasing the conditional
probability that links the emergence of one event to the emergence of previous events (Cover & Thomas, 2006).

For instance, the conditional probability between an actual left target (trial n) and a previous right target (trial n-1) is 1 if a left always follows a right. In this specific deterministic case, conditional probabilities can be assigned for each possible combination of previous and present events to create patterns, i.e. right after left and left after right for a LRLRLR pattern. In this way, the use of Markov chains allows the creation of different patterns (regularities) in a specific event sequence. This concept can be extrapolated by assigning conditional probabilities to events that are several steps back in the event sequence. In this way, it is possible to postulate values of conditional probabilities between the present and two previous events (trial n-2), three previous events (trial n-3) and so on (trial n-X; Meyn & Tweedie, 2008).

In a binary decision making task with two choices, conditional probabilities of 0.5 mean that there is no correlational link between the actual and the previous event (random sequence). In that case, the event sequence is called an order 0 Markov chain. An event sequence is classified as an order 1 Markov chain, if the conditional probabilities between the actual trial (n) and the previous trial (n-1) are higher or lower than 0.5. An event sequence is classified as an order 2 Markov chain, if the conditional probabilities are higher or lower than 0.5 and link the actual trial (n) with the two previous trials (n-1 and n-2). This way, higher order Markov chains can be created (Cover & Thomas, 2006). As the order of the Markov chain increases, more complex patterns can be generated. More specifically, participants only need to remember the previous trial for an order 1 Markov chain but they need to remember the two previous trials in an order 2 Markov chain, and so on. This increases the difficulty to find the pattern in the sequence and to learn the optimal strategy for the task. Indeed, the higher the order of the Markov chain sequence, the more the participant needs to explore different hypothesis about how the sequence works at the beginning of the task in order to succeed in the learning task. Hence, they need to ignore the exploitation of simpler strategies they discovered at the beginning of the task (see Sutton & Barto, 1998).

For this study, three Markov sequences were designed: an order 0 Markov sequence was used in block 1 and two distinct order 2 Markov sequences were used in blocks 2 and 3. More details are described below.
2.5.3.3.2. Target option probabilities and conditional probabilities in block 1

In block 1, the target option probability for one option was always different from the other option, i.e. 0.72 vs. 0.28. The conditional probability of 0.5 represented a random event sequence. Therefore, the optimal learning strategy for block 1 was the repeated choice of the target position with the highest target probability resulting in a maximum accuracy of 72%. This strategy is known as maximization (Vulkan, 2000).

2.5.3.3.3. Target option probabilities and conditional probabilities in block 2

In block 2, the target option probability was 0.5, meaning that the likelihood of winning was the same for both target positions and at chance level. Hence, participants could not repeat their block 1 learning strategy to succeed in this block 2. However, the event sequence of block 2 had an order 2 Markov probabilistic chain, presenting a hidden trial pattern sequence in the block (LLRRLRRLRLL…). The conditional probabilities linking the events two previous trials ago (n-2) to the actual trial (n) are shown in Table 2.1.
Table 2.1

Conditional probabilities between events two previous trials ago (n-2) and the actual trial in the event sequences employed in block 2

<table>
<thead>
<tr>
<th>Two previous trials (n-2)</th>
<th>Actual trial (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Left</td>
</tr>
<tr>
<td>Left/Left</td>
<td>0.27</td>
</tr>
<tr>
<td>Left/Right</td>
<td>0</td>
</tr>
<tr>
<td>Right/Left</td>
<td>1</td>
</tr>
<tr>
<td>Right/Right</td>
<td>0.73</td>
</tr>
</tbody>
</table>

In this event sequence, the trial pattern sequence is disrupted at times because conditional probabilities linking the trials are different than 1 (e.g. LLLRRLLRRLLRRR…). Therefore, the learning of the pattern was difficult and uncertain. If the participant was able to learn the pattern and disregard the eventual disruptions, the maximum accuracy in this block was approximately 82%. Note, the conditional probabilities used in block 2 were very similar to the target option probabilities used in block 1 (option A: 0.72; option B: 0.28). In this way, the disruptions were designed to be similar between blocks.

The pattern was always the same throughout the block but the disruption in the pattern sequence means that participants have to constantly monitor and re-assess their internal models about the optimal strategy in this block and to decide to continue with the exploitation of the detected pattern or to explore the environment for new learning strategies. This reassessment of strategies is especially evident at the beginning of the block but it should still occur throughout the block. In contrast, the dilemma of the usage of exploration versus exploitation strategies ceases to exist in a deterministic pattern sequence as soon as participants have learnt the pattern.

2.5.3.3.4. Target option probabilities and conditional probabilities in block 3

Block 3 is a combination of both block 1 and block 2 strategies. First, probabilities between options are 0.72/0.28, as it was for block 1. Secondly, conditional probabilities were not random here, as it was for block 2. In this way, an order 2 Markov probabilistic model, presenting a hidden pattern sequence in the block was present (RRLRRLRR…). The conditional probabilities linking two previous trials to one actual trial at block 3 are shown at table 2.2.
Table 2.2

*Conditional probabilities between two previous trials and one actual trial for the sequences employed at block 3*

<table>
<thead>
<tr>
<th>Two previous trials (n-2)</th>
<th>Actual trial (n)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Left</td>
<td>Right</td>
<td></td>
</tr>
<tr>
<td>Left/Left</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Left/Right</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Right/Left</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Right/Right</td>
<td>0.6</td>
<td>0.4</td>
<td></td>
</tr>
</tbody>
</table>

As it is possible to see at table 2.2, this pattern is different from the one employed at block 2, so participants would need to learn the pattern again. Additionally, even though probabilities between options were 0.72/0.28 as it was for block 1, at block 3, the option presenting the most of targets was the opposite one as block’s 1 option (i.e. left for block 1 and right for block 3). In this way, participants would also need to learn which side presented more targets again.

This trial pattern sequence is disrupted at times (e.g. LLRLRRLLRLR…), as at block 2. Participants can choose to maximize in one side or to follow a specific pattern in the sequence. If they choose maximization, the maximum accuracy is 72%, while if they choose the pattern, disregarding the eventual disruptions, the maximum accuracy would be 82%. In this way, the pattern strategy is more efficient.

At first, the design aimed to create sequences with the same maximum accuracy for blocks 1, 2 and the two strategies at block 3. This study is shown at pilot 1 (Chapter 4). Unfortunately, the combination of the same maximum accuracies between blocks, left us with a maximum accuracy value of 0.66 for the patterns. Participants were not able to learn the pattern strategy at all, so it was decided to increase maximum accuracy values for blocks 2 and 3 patterns. At the end, sequences with a maximum accuracy of 0.82 were selected for blocks 2 and 3, so at least, two blocks presented the same maximum accuracy. A maximum accuracy of 0.82 for block 1 was not considered, as this would be too easy for participants. At the end, it was decided to keep blocks 2 and 3 with maximum accuracies of 0.82 and 0.72 for block 1. However, at block 3, the maximum accuracy of 0.82 for the pattern strategy was only possible if a distinct maximum accuracy was present for the between option strategy.
This created an environment where one strategy presented a higher maximum accuracy (pattern) than the other strategy (between options).

2.5.3.3.5. **Controlling for possible confounding variables related to the block design**

The block order was always the same for all experiments. Block 1 (target option block) was presented first, so this avoided participants being exposed to previous pattern search strategies of the other blocks. In fact, always putting this block at first decreased the probability to increase the bias in participants to search for patterns in block 1, as it was shown that participants generally look for patterns, even when the sequence is completely random (Vulkan, 2000; Unturbe & Corominas, 2007). Block 2 (pattern search block) was presented next to avoid an immediate repetition of block 1 strategy in the subsequent block. Block 3 (target option and pattern search block) was presented last because with this block order participants could be exposed to one strategy at a time before combining both strategies in the final block.

Individually distinct event sequences were created for each participant and blocks. These individual sequences were matched for the checker and non-checker groups, i.e. the first checking participant used the same sequences as the first non-checking participant. The target option and conditional probabilities of these individual sequences for blocks 1, 2, and 3 were the same as specified above. These individual event sequences were created to exclude the possibility that one specific event sequence structure would cause the reported effects.

Finally, the target location with the higher target probability was counterbalanced across participants in block 1, i.e. half the participants had the target with the higher target probability on the left side and the other half of participants had it on the right side. Moreover, the target location with the higher target probability was always the alternative target option in block 3 relative to block 1 (e.g. block 1: left target / block 3: right target).

2.5.3.4. **Binary decision making task payment**

The task of the participant was to win as much money as possible (positive-neutral and positive-positive task) or to lose as little money as possible (negative-neutral, negative-negative or symptom-related experiment) at the end of the Binary decision making task. Participants could win/lose 2 pence per trial. The amount of 2 pence per trial was calculated based on an average accuracy of young adults in these types of experiments (approx. 60%; Vulkan, 2000). Based on this calculations, participants won approx. £6 in positive monetary
feedback experiments, and lost approx. £4 of the gambling money (£10) in the negative monetary feedback experiments. Therefore, all participants received about £16 (£10 for participation and £6 winnings from the Binary decision making task) at UoS and £18 (£12 for participation and £6 winnings from the Binary decision making task) at BU, irrespectively of the experimental feedback profile of the task.

2.5.4. Laboratory-based questionnaires

At the end of the laboratory-based session, participants were asked to fill out four questionnaires on a computer (programmed in E-Prime v2.0) which were the Domain Specific Risk-Taking Scale (DOSPERT; Blais & Weber, 2006), the Intolerance of Uncertainty Scale (IUS; Buhr & Dugas, 2002); the UPPS Impulsivity Scale (Whitehead & Lynam, 2001), and the South Oaks Gambling Screen (Lesieur & Blume, 1987). In addition, they filled out a paper-based strategy questionnaire (Appendix 2.6).

All the questionnaires are standard research tools, except for the strategy questionnaire which was specifically designed for this study. The questionnaires were used to evaluate potentially correlating self-reported measures such as intolerance to uncertainty, risk taking behaviour, and impulsivity. All these are common cognitions and behaviours associated with OCD (Tolin, Abramowitz, Brigidi & Foa, 2003 for IU; Frost, Steketee, Cohn & Griess, 1994 for risk taking; Summerfeldt, Hood, Antony, Richter & Swinson, 2004 for impulsivity). Specific details for each questionnaire are described below.

2.5.4.1. Detailed laboratory-based questionnaire descriptions

2.5.4.1.1. The Domain Specific Risk-Taking Scale (DOSPERT)

The DOSPERT questionnaire is a psychometric scale that assesses risk-taking in five domains: finance, health/safety, recreational, ethical and social decisions (Blais & Weber, 2006). The DOSPERT total and finance scores for risk-taking and risk-perception were the main interest of this study. Examples for each domain are: “investing 10% of your annual income in a new business venture” (finance); “engaging in unprotected sex” (health/safety); “taking a weekend sky-diving class” (recreational); “Having an affair with a married man/woman” (ethical); “disagreeing with an authority figure on a major issue” (social); (Blais & Weber, 2006). The questionnaire contains 30 items (6 items relating to each domain) and comprises two sections: risk-taking (how likely someone engages in that behaviour) and
risk-perception (how risky each behaviour is perceived to be). Each question is rated on a 7-point Likert scale which evaluates the likelihood the respondents might engage in risky behaviours (Blais & Weber, 2006). The higher the scores, the greater the risk taking and perception of risk in each domain. The domains have good internal reliability ($\alpha = .78$ and $\alpha = .77$ for the risk taking scales and risk perception scales, respectively) a good test-retest reliability, as well as a good construct validity reported by Weber, Blais & Betz (2002). The DOSPERT finance and total scores were recorded to assess potential correlations between risk-taking and reward-based decision making and learning measures from the Binary decision making task.

2.5.4.1.2. Intolerance of uncertainty questionnaire (IUS)

The IUS is a self-report measure that assesses beliefs that (a) uncertainty is stressful and upsetting, (b) uncertainty leads to the inability to act, (c) uncertain events are negative and should be avoided, and (d) being uncertain is unfair (Buhr & Dugas, 2002). This questionnaire has two subscales: Factor 1 means uncertainty has negative behavioural and self-referent applications and Factor 2 means uncertainty is unfair and spoils everything. The IUS has an excellent internal consistency ($\alpha = 0.94$) and a good test-retest reliability (Buhr & Dugas, 2002; Norton, 2005). This questionnaire is a 27-item scale with 5-point Likert scale answers from 1 (“not at all characteristic of me”) to 5 (“entirely characteristic of me”). Higher total scores indicate higher levels of intolerance of uncertainty. Scores under 40 reflect some intolerance to uncertainty, scores of $\geq 50$ reflect some problems with uncertainty, and scores of $\geq 70$ suggest intolerance of uncertainty (Leahy, Holland, & McGinn, 2011). This questionnaire was employed to investigate possible correlations between intolerance of uncertainty scores and reward-based decision making and learning behaviour in uncertain environments, especially because there is evidence OCD present a higher intolerance of uncertainty (Holaway et al., 2006; see also Chapter 1.5.1).

2.5.4.1.3. The UPPS Impulsive Behaviour Scale

The UPPS questionnaire presents 45 questions and was designed to measure impulsivity across 4 distinct dimensions, rating subtypes of urgency, sensation seeking; lack of premeditation and lack of perseverance. Urgency is rated by 12 items, with higher scores meaning individuals are likely to engage in impulsive behaviours, acting less quickly, particularly in the face of negative effect. Sensation-seeking is rated by 12 items, higher
scorers meaning participants enjoy taking risks and engaging in dangerous activities. Lack of premeditation is rated by 11 items and high scorers are likely to act on the spur of the moment. Finally lack of perseverance is rated by 10 items and higher scores indicate the individual is not capable of working in conditions which require resistance to distracting stimuli. Each question is rated on a 4 point Likert scale with 1 = “strongly agree” and 4 = “disagree strongly”. Internal consistency for the four factors was $\alpha = .89$ for urgency; $\alpha = .87$ for lack of premeditation; $\alpha = .88$ for lack of perseverance and $\alpha = .90$ for sensation seeking (Whiteside & Linam, 2003).

As impulsivity was associated with OCD (Summerfeldt et al., 2004), this questionnaire was employed to investigate possible correlations between the UPPS scores and reward-based decision making and learning in the Binary decision making task.

### 2.5.4.1.4. South-Oaks Gambling Screen

The South Oaks Gambling Screen is a 20-item questionnaire based on DSM-III criteria for pathological gambling. It offers a means to screen the general population, regarding pathological gambling. Its test-retest correlation was $r = .71$ and its internal consistency was $\alpha = .97$ (Lesieur & Blume, 1997). Each question is rated on the basis of the type of gambling the participant have done in her lifetime and participants should answer “yes” or “no” for some questions or tick an alternative in a multiple choice question. Each “yes” represents one point. The sum of these points is used for further analysis. Zero points mean “no problem with gambling”; 1-4 points mean “some problem with gambling”, and 5+ points mean “probable pathological gambler”. Please note, the multiple choice questions are not counted into the sum score, as they do not relate to the risk of pathological gambling. This questionnaire was employed to exclude possible pathological gamblers from the population to be studied, as this could affect results on the Binary decision making task.

### 2.5.4.1.5. Strategy questionnaire

A 6-item questionnaire recorded the decision making strategies that participants used during the experiment. This short questionnaire was especially designed for this study. It aimed to evaluate whether participants were consciously using strategies for blocks 1, 2 and 3 (yes vs. no answer; categorical data), and whether they were consciously visualizing a specific pattern structure for blocks 1, 2 and 3 (yes vs. no answer; categorical data). This
questionnaire was designed to evaluate group differences in the explicit knowledge of block strategies.

2.6. Data analysis of the laboratory-based part of the study

2.6.1. Tests of Normality

Averages and standard errors were calculated and tests of normality were conducted by calculating skewness and kurtosis z-scores for all variables. Z-scores ≥ 1.96 were considered as violations of the skewness and kurtosis normality assumptions (Field, 2009). Non-parametric tests were performed on data with violations of the normality assumptions (e.g. Mann-Whitney U tests, Wilcoxon tests etc) wherever possible.

Note, we continued to use ANOVAs where necessary even if some variables were non-parametric because of the absence of an equivalent non-parametric analysis that could be applied without much loss of data information. Levene tests were performed for mixed and between-participant ANOVAs to test for the homogeneity of variance assumption. Repeated measures violations of the sphericity assumption were corrected using the Huynh-Feldt correction of the degrees of freedom of the F-distribution. Post-hoc ANOVAs and t-tests were conducted and Bonferroni corrections were applied when necessary.

2.6.2. Data analysis for the Picture Validation Task

In Chapter 3, valence, arousal and comfort ratings for positive, neutral, negative and checker-related pictures were analysed in three separate mixed 2 x 4 ANOVAs with the between-participant factor Group (checkers vs. non-checkers) and the within-participant factor Picture Type (positive, negative, neutral vs. checking-related pictures). In Chapters 5-7, the between-participant factor Experiment was added to the mixed ANOVAs.

2.6.3. Data Analysis for the Binary Decision Making Task

2.6.3.1. Accuracy measures

Accuracy is a widely used measure in reward based decision making and learning experiments (Vulkan, 2000). It reflects the performance level in the binary decision making task. Statistical analyses of accuracy measures were conducted across blocks and separately for each block because blocks had different optimal strategies.
In Chapter 5, a two-way 2 x 2 ANOVA with the between-participant factors Group (checkers vs. non-checkers) and Experiment (e.g. positive-neutral experiment vs. negative-neutral experiment) was conducted for each block to investigate accuracy levels in the three blocks across experiments and groups.

In Chapter 6, a two-way 2 x 2 between-participant ANOVA with the factors Group (checkers vs. non-checkers) and Experiment (positive-neutral vs. positive-positive experiment) was computed for each block to investigate the effect of the reward magnitude enhancement due to the presentation of positive pictures after the positive feedback when comparing checkers and non-checkers. A second between-participant two-way ANOVA with the factors Group (checkers vs. non-checkers) and Experiment (negative-neutral vs. negative-negative experiment) was computed for each block to investigate the effect of the reward magnitude enhancement due to the presentation of negative pictures after the negative feedback comparing checkers and non-checkers.

In Chapter 7, a two-way between-participant ANOVA with the factors Group (checkers vs. non-checkers) and Experiment (negative-neutral vs. symptom related experiment) was computed for each block to investigate the disease related reward magnitude alteration in checkers compared to non-checkers when using symptom-related vs. neutral pictures.

In Chapters 5-7, additional accuracy analyses were conducted to investigate how participants were learning over the course of a block. For this analysis, each block was divided in 5 sub-blocks and analysis were conducted for each block. The within-participant factor sub-blocks (1, 2, 3, 4 vs. 5) was added to the analyses and a three-way mixed ANOVA with the between-participant factors Group and Experiment was executed.

2.6.3.2. **Exploration / exploitation strategy analysis: probability of shifting after winning in the previous trial (win/shift probability) and probability of shifting after losing in the previous trial (lose/shift probability)**

This analysis was conducted to investigate the effect of winning or losing in the previous trial on the decisions in the current trial, i.e. stay with current option or shift to other option. This analysis was only conducted on the data of block 1 because the shifting probabilities in blocks 2 and 3 were not only dependent from previous payoff but also on the pattern presented in the experimental sequences. Here, the dependent variables win/shift probability and lose/shift probability were analysed, separately. A two-way 2 x 2 ANOVA
with the between-participant factors Group (checkers vs. non-checkers) and Experiment (e.g. positive-neutral experiment vs. negative-neutral experiment) was conducted for win/shift and lose/shift separately. We did not perform any sub-block analyses for these dependent variables.

2.6.3.3. Cross-correlations and auto-correlations

2.6.3.3.1. Cross-correlations between the stimulus and the response sequence

This cross-correlation analysis was conducted to investigate whether participants are using information from previous trials to make decisions in the actual trial (examination of sequential effects). This analysis was especially relevant for blocks 2 and 3 because it can show pattern learning effects.

Cross-correlations are the probability of correlations between a participant’s response in the actual trial and the preceding events in the experimental sequence. These correlations were calculated for the actual response and the immediately preceding event (lag 1, n-1) and for the actual response and all preceding events until lag 5 (lag 5, n-5). Positive correlations indicate an increased probability of choosing the same option as the stimulus presented in a preceding trial at lag x, whereas negative correlations are caused by an increased likelihood of choosing the alternative option compared to lag x.

Therefore, cross-correlations between the current trial (n) and previous trials (e.g. lag 1 (n-1), lag 2 (n-2)) can indicate whether participants successfully try to infer complex regularities between trials. For example, in a “RRL” pattern sequence, the target is present on the right side two trials before (n-2) and on the left side one trial before (n-1) the actual event. In this case, participants have to keep several trials in their working memory to be able to follow a stimulus pattern with their responses. Their behaviour is based on perceived sequence regularities.

Please note, the win/shift and lose/shift probabilities that are calculated in relation to the immediately preceding trials (n-1) are a special case of cross-correlation. They do not only give information about stay or shift correlations but also take into account wins and losses in the previous trial. Cross-correlation calculations were based on a formula used by DeGroot & Schervish, (2012) (Figure 2.5).
2.5 Adapted formula from DeGroot & Schervish, (2012) to make the calculus of cross-correlation and auto-correlation.

\[ C_{xy}(r) = \frac{1}{N-r} \sum_{i=1}^{N-r} x_i y_i + r \]

Based on this formula, a moving window analysis was conducted to correlate the response and stimulus sequences. More specifically, if there was a match between the participant’s response type (left vs. right) in the current trial and the stimulus type (left vs. right) at a specific lag (e.g. lag 1; previous trial) the value 1 was given. When there was no match, the value -1 was given. The sum of these values divided by the total number of values resulted in the cross-correlation probability for a specific lag. The same analysis was repeated for lags 1 to 5 for each participant and block. Finally, means and standard errors were calculated for each group, block, and experiment.

2.6.3.3.2. Stimulus sequence auto-correlations

The auto-correlation analysis was very similar to the cross-correlation analysis. The only difference was that the correlations are calculated by comparing the experimental stimulus sequence in the current trial (n) to the experimental stimulus sequences for the previous trials (lags 1-5; n-1 to n-5). For example, the auto-correlation of a pattern sequence like LLRRLLL would show higher correlation values of opposite polarity at n-2 and n-4 because the sequence repeats itself after four trials and the stimulus type changes after two trials which is mirrored in the polarity reversal of the correlation (Figure 2.6). Please note, there is no pattern in the auto-correlations when the sequence is random, as in block 1, but an auto-correlation pattern can be found for the stimulus sequences used in blocks 2 and 3.

Disruptions of event sequences with patterns by added noise are reflected in the auto-correlations. For example, the auto-correlation for a LLRRLLL pattern without disruptions (no noise) is maximal at lags 2 and 4, as shown at figure 2.6 (blue line). However, if disruptions are added to this pattern sequence, correlation value magnitudes reduce, depending on the amount of disruptions inserted for the sequence. Figure 2.6 shows the same pattern sequence with disruptions (red line).
Figure 2.6 Auto-correlation for a hypothetical pattern sequence (LLRRLL…) for lags 1-5. Blue line: Auto-correlation for the pattern sequence without disruptions (no noise). Red line: Auto-correlation for the pattern sequence with disruptions (added noise).

In summary, auto-correlations contain information about the conditional probabilities asserted for an experimental sequence. For example, if a right stimulus would always follow a right stimulus in a sequence (no disruptions), the lag 1 (n-1) autocorrelation would be always 1. If that would be only the case in 80% of all trials (with disruptions), the conditional probability value would decrease. Auto-correlations also reveal the pattern of a sequence. In this way, it is possible to use sequence’s auto-correlation analysis to investigate how the sequence is generating patterns and how the disruptions are affecting this pattern. This was important for the selection of the best sequences for this study (Chapter 4).

2.6.3.3. Statistical comparison of auto-and cross-correlations

In the same way that auto-correlation of the sequence informs about the amount of disruptions in the pattern sequence, the comparison between the auto-correlation of a stimulus sequence and the cross-correlation between the stimulus sequence and the participant’s response sequence can reveal how close the response pattern of the participant was to the experimental pattern sequence (Figure 2.7).

For this thesis, auto-correlation and cross-correlations were calculated for each participant and block sequences for the lags 1-5 (n-1 to n-5). Afterwards, mean differences and related standard errors between these auto- and cross-correlations (AC\text{value} - CC\text{value}) were used to compare groups, blocks and experiments. The closer to zero this difference, the more similar the cross-correlation and the auto-correlation values are for a particular lag. In this way, mean difference values indicated how many disruptions from the original pattern are reflected in the participant’s response sequence performance (Figure 2.7).
Figure 2.7 Hypothetical values for the auto-correlation of a sequence (experimental sequence) and of the cross-correlations for two groups, calculated for lags 1-5, indicating a pattern search strategy. The auto-correlation represents a LLRR pattern sequence with disruptions. Group 1 represents a cross-correlation with similar values relative to the auto-correlation values, indicating pattern searching. The cross-correlation for Group 2 is less similar to the auto-correlation values, indicating that this group learned the patterns less well than Group 1, as there were more disruptions in the pattern of their response sequence. Please note, the mean difference between group 1 and the auto-correlation of the experimental sequence is closer to zero than the mean difference for group 2.

Additionally, the comparison between the stimulus sequence and the participant’s response sequence can reveal if the participant is using a different decision making strategy that is distinct from pattern searching, for example maximization. For instance, the level of similarities between the cross-correlation and auto-correlation values for lag 1 to 5 reflects the performance level of the participant for this sequence in the task and reveals the strategy the participant is using. If the original sequence presents a pattern and participants present a response sequence that is similar to the original sequence, this analysis reveals that the participants is searching for patterns in the sequence (Figure 2.8a). However, if the participant response sequence does not present cross-correlation peaks where the pattern usually repeats itself, this could indicate others strategies. For instance, if a participant performs maximization and repeats the same response during all the task, the participant response sequence will present a huge amount of repetition (i.e., LLLLLLLLLLLL). If the response sequence presents more targets in this same option (i.e., LLLLLLLLLRLR), all the
time the participant response sequence presents a response in the option with most targets (maximization), the response sequence will correlate with these responses, increasing the probability of cross-correlation from lags 1 to 5 between these two sequences. Indeed, as the cross-correlation probabilities will increase from lags 1 to 5, the maximization strategy can be revealed when the cross-correlation values from lags 1 to 5 are similar to each other, revealing a line in the figure (figure 2.8). In this way, differently from a pattern searcher, where cross-correlation peaks are present, a maximizer will present similar values of cross-correlation forming as a line (figure 2.8).

![Figure 2.8 Hypothetical values for the auto-correlation of a sequence (experimental sequence) and the cross-correlations for two groups, calculated for lags 1-5, indicating a pattern searcher strategy and a maximize strategy. The auto-correlation represents a LLR pattern sequence with disruptions. The pattern searcher line represents a cross-correlation with similar values relative to the auto-correlation values, indicating pattern searching. The cross-correlation for the maximizer is less similar to the auto-correlation values and indicates this group is repeating the same response throughout the task, increasing the cross-correlation values for all lags, indication maximization.](image)

In Chapter 5, a three-way mixed ANOVA with the within-participant factor Lag (1-5) and the between-participant factors Group (checkers vs. non-checkers) and Experiment (e.g. positive-neutral experiment vs. negative-neutral experiment) were conducted on the difference values to investigate monetary feedback effects on checkers and non-checkers.
In Chapter 6, a three-way mixed ANOVA with the within-participant factor Lag (1-5) and the between-participant factors Group (checkers vs. non-checkers) and Experiment (positive-neutral vs. positive-positive experiment) was computed on the difference values to investigate the effect of the reward magnitude enhancement by presenting positive pictures after the positive feedback on checkers and non-checkers. A second three-way mixed ANOVA with the within-participant factor Lag (1-5) and the between-participant factors Group (checkers vs. non-checkers) and Experiment (negative-neutral vs. negative-negative experiment) was computed on the difference values to investigate the effect of the penalty magnitude enhancement by presenting negative pictures after the negative feedback on checkers and non-checkers.

In Chapter 7, a three-way mixed ANOVA with the within-participant factor Lag (1-5) and the between-participant factors Group (checkers vs. non-checkers) and Experiment (negative-neutral vs. symptom related experiment) was computed on the difference values to investigate the effect of the disease-related emotional information on feedback magnitude alterations.

### 2.6.4. Data analysis of the laboratory-based questionnaires.

In Chapter 5, a two-way between-participant ANOVA with the factors Group (checkers vs. non-checkers) and Experiment (e.g. positive-neutral experiment vs. negative-neutral experiment) was conducted for each questionnaire in order to investigate questionnaire responses variations on checkers and non-checkers between groups and experiments.

In Chapter 6, a two-way between-participant ANOVA with the factors Group (checkers vs. non-checkers) and Experiment (positive-neutral vs. positive-positive experiment) was computed for each lab-based questionnaire to investigate the effect of these measurements when comparing checkers and non-checkers. A second between-participant two-way ANOVA with the factors Group (checkers vs. non-checkers) and Experiment (negative-neutral vs. negative-negative experiment) was computed.

In Chapter 7, a between-participant two-way ANOVA with the factors Group (checkers vs. non-checkers) and Experiment (negative-neutral vs. symptom related experiment) was computed to investigate these questionnaire measurements in checkers and non-checkers across experiments.
In Chapters 5-7, the strategy questionnaire data were analysed with Pearson Chi-square tests or Fisher’s exact tests to compare groups, separately for each experiment.

### 2.6.5. Correlations between questionnaire scores and behavioural measures from the binary decision making task

Kendall’s Tau correlations were calculated to investigate relationships between the self-report questionnaire measures and the key behavioural measures from the binary decision making task for the experiments reported in Chapters 5-7. The key behavioural measures from the binary decision making task were accuracy, win-shift, lose-shift, cross-correlations for block 1 (lags 1-5), block 2 (lags 1-5) and block 3 (lags 1-5). Kendall’s Tau correlations are more robust for small sample sizes (<30 participants). Significant correlations were considered weak for r-scores between 0.20 - 0.39; moderate for r-scores between 0.40 - 0.59 and strong for r-scores above 0.59 (Evans, 1996).

The correlation analysis was conducted for checkers to answer the following questions: 1) Is disease severity (OCI-R total and OCI-R checking scores) in checkers related to the performance in the binary decision making task? 2) Can correlations between the comorbidities depression and anxiety measured by the DASS and behavioural measures in the binary decision making task partly explain our findings in checkers? and 3) Do the laboratory-based measures for intolerance of uncertainty, financial (and total) risk taking and risk perception, impulsivity and gambling behaviour correlate with the behavioural measures of the binary decision making task in checkers?
Chapter 3: The Picture Validation Task

3.1. Introduction

The aim of the study presented in this chapter was to validate the emotional picture ratings of 100 positive, neutral and negative pictures from the International Affective Picture System (IAPS) database (Lang et al., 2008) and 25 checking-related from the OCD picture data base published by Simon et al. (2010). This validation study (study 1) was conducted, firstly, to ensure that representative samples of different picture categories were selected for the future usage of these pictures in the binary decision making tasks described in Chapters 4-7. Secondly, to evaluate the valence level of the checking-related pictures in relation to the pictures selected from IAPS. Finally, to investigate picture rating differences between a subclinical checker group and a non-checker group. This would be an extension of the findings by Simon et al. (2010). They investigated OCD patients and a control group. In this validation study, subclinical checkers were expected to rate checking-related pictures more negatively than non-checkers. We did not expect any group difference in picture ratings for the positive, neutral and negative pictures.

The picture validation task used the same valence and arousal rating scales as the Self-Assessment Manikin (SAM) questionnaire by Bradley & Lang (1994) in their IAPS picture validation study. These rating scales were selected because emotional responses have two independent dimensions which are the physiological response and the cognitive label, as described in the two-factor model by Schachter & Singer (1962). Valence and arousal ratings are associated with these two independent dimensions. Valence ratings are related to the cognitive dimension of an experienced emotion (pleasant / unpleasant) whereas arousal ratings are a self-description of the physiological state of the body (calm / exciting).

As a third scale, the IAPS uses a dominance rating scale (controlled / in control) that reflects the amount of overt behavioural acts, associated to an object (Bradley & Lang, 1994). However, to the purpose of this thesis, it was decided to exchange the IAPS dominance rating scale with a comfort rating scale (comfortable / uncomfortable). This was done because studies have shown that compulsions are commonly used to decrease discomfort caused by obsessions in OCD patients (Veale et a., 1993; Veale, 2007 for a review). In this way, it seemed a new comfort scale could facilitate the emotional rating of checker participants, as possible obsessions triggered by the pictures could be easily reflected by a comfort level rating.
Additionally, it was not one of the aims of this study to assess behavioural acts (dominance) to emotional pictures. The main focus was related with the experience of emotions.

Hence, valence, arousal and comfort ratings were measured in the picture validation task. Based on these initial picture ratings, 100 pictures (40 neutral, 20 positive, 20 negative, and 20 checking-related pictures) were selected for the binary decision making studies. This selection was based on two criteria. The first criterion was to exclude pictures that had the smallest mean valence rating difference between checkers and non-checkers for checking-related pictures and the largest mean valence rating difference between checkers and non-checkers for positive, negative, and neutral pictures. The second criterion was to exclude pictures originally classified as neutral but rated as positive or negative emotions in the current sample.

These criteria were chosen because we aimed to trigger the same amount of valence-related reinforcement in the binary decision making tasks for both groups for positive, neutral, and negative pictures. The exception was the symptom-related experiment where the picture type was expected to differently influence reward-based learning and decision making processes in checkers and non-checkers by altering the feedback magnitude differentially.

3.2. Methods

3.2.1. Participants

The online questionnaire was completed by 69 participants recruited from the University of Surrey. Based on the questionnaire responses, 11 checkers and 11 non-checkers were recruited for the laboratory based picture validation task. The reasons for exclusion from the laboratory-based study are shown in Table 3.1.
Table 3.1.

*Reasons for exclusion from the laboratory based study*

<table>
<thead>
<tr>
<th>Exclusion criterion</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incomplete online questionnaire</td>
<td>16</td>
</tr>
<tr>
<td>Does not want to take part in PVT</td>
<td>3</td>
</tr>
<tr>
<td>Age above exclusion criteria</td>
<td>1</td>
</tr>
<tr>
<td>Diagnosed Depression</td>
<td>5</td>
</tr>
<tr>
<td>DASS score above criteria</td>
<td>6</td>
</tr>
<tr>
<td>Neither checker nor non-checker</td>
<td>6</td>
</tr>
<tr>
<td>Gave up after being selected</td>
<td>10</td>
</tr>
</tbody>
</table>

As shown in Table 3.2, the checker and non-checker groups in the laboratory-based study were matched for age and gender. Checkers had significantly higher OCI-R checking subscale scores compared to non-checkers, $t(20) = 12.35, p < 0.001$, which was expected due to the selection criteria. In addition, checkers also had significantly higher scores for the OCI-R total and the OCI-R ordering subscale and the OCI-R counting subscales (all $p < .001$). In addition, the DASS anxiety, depression, and stress scores were recorded to measure potential comorbidities. As expected, checkers had a significantly higher DASS anxiety score than non-checkers, $U=17, z = -2.88, p = .003$. They were more anxious. There were no group differences for depression and stress. Furthermore, the videogame experience and the habit to play for money did not differ significantly between groups.
Table 3.2

Checker and non-checker group similarities and differences for the OCI-R scores and DASS scores. \( F = \) female, \( M = \) male, \( Pv = \) videogame player, \( NPv = \) no videogame player, \( Pm = \) money player, \( NPM = \) no money player

<table>
<thead>
<tr>
<th>Questionnaires</th>
<th>Checkers</th>
<th>Non-Checkers</th>
<th>Statistical Test</th>
<th>Effect sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographic Questionnaire</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>22 ± 0.77</td>
<td>22.36 ± 1.01</td>
<td>( t(20) = 0.28, p = .77 )</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>10 F, 1 M</td>
<td>9 F, 2 M</td>
<td>( p = 1 )</td>
<td></td>
</tr>
<tr>
<td>Play videogame</td>
<td>4 Pv, 7 NPv</td>
<td>3 Pv, 8 NPv</td>
<td>( p = 1 )</td>
<td></td>
</tr>
<tr>
<td>Play for money</td>
<td>0 Pm, 11 NPm</td>
<td>0 Pm, 11 NPm</td>
<td>( p = 1 )</td>
<td></td>
</tr>
<tr>
<td><strong>OCI-R</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OCI-R checking</td>
<td>6.36 ± 0.33</td>
<td>0.91 ± 0.28</td>
<td>( t(20) = 12.35, p &lt; .001^{**} ) ( r = .94 )</td>
<td></td>
</tr>
<tr>
<td>OCI-R washing</td>
<td>3.00 ± 1.03</td>
<td>0.82 ± 0.27</td>
<td>( t(20) = 2.04, p = .05^{*} )</td>
<td></td>
</tr>
<tr>
<td>OCI-R ordering</td>
<td>6.00 ± 0.89</td>
<td>1.09 ± 0.28</td>
<td>( t(12.00) = 5.23, p &lt; .001^{**} ) ( r = .67 )</td>
<td></td>
</tr>
<tr>
<td>OCI-R counting</td>
<td>3.09 ± 0.79</td>
<td>0.27 ± 0.19</td>
<td>( U = 15.50, z = -2.16, p = .002^{**} ) ( r = .61 )</td>
<td></td>
</tr>
<tr>
<td>OCI-R obsessing</td>
<td>3.55 ± 1.01</td>
<td>1.91 ± 0.68</td>
<td>( t(17.50) = 1.34, p = .19 )</td>
<td></td>
</tr>
<tr>
<td>OCI-R total</td>
<td>27.18 ± 3.21</td>
<td>7.64 ± 1.05</td>
<td>( t(12.13) = 5.77, p = .001^{**} ) ( r = .79 )</td>
<td></td>
</tr>
<tr>
<td><strong>DASS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DASS depression</td>
<td>3.18 ± 1.19</td>
<td>6.73 ± 1.71</td>
<td>( t(17.19) = 1.73, p = .10 )</td>
<td></td>
</tr>
<tr>
<td>DASS anxiety</td>
<td>7.18 ± 0.93</td>
<td>2.36 ± 0.93</td>
<td>( U = 17, z = -2.88, p = .003^{**} ) ( r = .61 )</td>
<td></td>
</tr>
<tr>
<td>DASS stress</td>
<td>10.91 ± 1.84</td>
<td>6.91 ± 1.46</td>
<td>( t(20) = 1.70, p = .10 )</td>
<td></td>
</tr>
</tbody>
</table>

3.2.2. Procedure and Materials

This study had two parts, an online screening questionnaire and a laboratory-based picture validation task (details in Chapter 2). For the picture validation task, pictures were rated on three rating scales, i.e. valence, arousal and comfort. More specifically, 25 positive, 50 neutral, 25 negative and 25 checking-related pictures were rated on each scale. Hence, the total...
trial number was 375. The number of neutral pictures was doubled because the binary decision making task experiments reported in Chapter 5 required the presentation of two sets of neutral pictures (see Chapter 5 for more information).

3.2.3. Data Analysis

The data analysis of the picture validation task is described in detail in Chapter 2. As a short reminder, picture ratings were analysed separately for valence, arousal and comfort values with a two-way 4 x 2 mixed ANOVA with the within-participant factor Picture Type (positive, neutral, negative and checking-related pictures) and the between-participant factor Group (checkers vs. non-checkers). Post-hoc t-tests were conducted and Bonferroni corrections applied when necessary.

After this analysis, the initially selected 125 pictures were reduced to 100 pictures which were used in the binary decision making experiments. The selection of pictures was based on two evaluation criteria (Table 3.3) based on the valence values because these show the differences between negative and positive emotions (Lang et al., 2008). Firstly, possible group differences between positive, neutral, and negative pictures were reduced by excluding the pictures with the largest group valence rating differences, i.e. exclusion of five positive pictures, 5 negative pictures, and of 10 neutral pictures. In contrast, the valence rating differences of the checking-related pictures were increased by excluding the five symptom-related pictures with the smallest group differences between checkers and non-checkers. For the neutral pictures, besides the first criterion, preselected neutral pictures that were actually rated as positive (<4) or negative (>6) were also excluded.
Table 3.3

Exclusion criteria for each picture type based on the valence scale ratings.

<table>
<thead>
<tr>
<th>Picture Type</th>
<th>Criteria for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Checking-related</td>
<td>Exclude five pictures with the smallest mean valence rating difference between checkers and non-checkers.</td>
</tr>
<tr>
<td>Negative</td>
<td>Exclude five pictures with the largest mean valence rating difference between checkers and non-checkers.</td>
</tr>
<tr>
<td>Positive</td>
<td>Exclude five pictures with the largest mean valence rating difference between checkers and non-checkers.</td>
</tr>
<tr>
<td>Neutral</td>
<td>Exclude pictures with the largest mean valence rating difference between checkers and non-checkers.</td>
</tr>
</tbody>
</table>

Evaluation criterion 2: Select pictures with valence ratings close to 5 (neutral)

<table>
<thead>
<tr>
<th>Picture Type</th>
<th>Criteria for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td>Exclude remaining pictures (total exclusion number = 10) when mean valence picture ratings &lt; 4 or &gt; 6 for checkers and non-checkers.</td>
</tr>
</tbody>
</table>

Besides this initial analysis, data from all the picture validation tasks performed in studies 2 to 6 (see Chapters 5 to 7) were posteriorly pooled altogether with data from this initial picture validation task study (study 1), as a way to investigate possible group differences between checkers and non-checkers for the positive, neutral, negative and checker-related picture categories for the valence, arousal and comfort scales, considering a bigger sample.

3.3. Results

3.3.1. Picture Validation task: Analysis of the selected pictures

3.3.1.1. Picture Valence Rating Analysis

Figure 3.1 shows the valence ratings for the selected picture types. As expected, positive pictures seem to have a lower valence compared to negative pictures, with intermediate scores for neutral pictures. Group valence ratings seem to not differ between
these picture types. Checkers seemed to have more negative ratings for checking-related pictures compared to non-checkers. Valence ratings of checking-related pictures seemed to be more negative than for neutral pictures but not as negative as for negative pictures.

Figure 3.1 mean valence ratings for positive, neutral, negative, and checking-related selected pictures displayed separately for checkers and non-checkers. Lower scores represent more positive valence and higher scores mean more negative valence. Neutral ratings have a score close to 5. Error bars represent standard error measures.

The ANOVA showed a significant main effect of Picture Type, $F(3, 60) = 302.51, p < .001$, and a significant interaction between the factors Picture Type and Group, $F(3, 60) = 7.23, p < .001$. Post-hoc paired sample t-tests showed that positive pictures ($2.21 \pm 0.14$) were rated more positively than neutral pictures ($4.48 \pm 0.15; t(21) = 10.92, p < .001$), and neutral pictures were rated more positively compared to negative pictures ($7.25 \pm 0.15; t(21) = 17.22, p < .001$). Negative pictures were rated as more negative than positive pictures, $t(21) = 22.37, p < .001$. Checking-related pictures were rated more positively than negative pictures ($5.68 \pm 0.17; t(21) = -10.61, p < .001$), and more negatively than neutral pictures ($t(21) = 7.03, p < .001$) and positive pictures ($t(21) = 14.00, p < .001$). Independent t-tests comparing picture types between groups did not reveal any group difference for the positive, neutral, and negative picture types. However, there was a significant difference for the checking-related pictures between checkers and non-checkers ((Mean (Checker): 6.25 ± 0.21; Mean (Non-checker): 5.12 ± 0.12; $t(15.63) = 4.48, p_b = .002, r = .75$).
3.3.1.2. Picture Arousal Rating Analysis

Figure 3.2 shows that arousal ratings varied between picture types. As expected, it seemed that positive pictures generated less arousal compared to negative pictures but similar arousal compared to checking-related pictures. In addition, the arousal levels for positive pictures were higher than for neutral pictures. For all picture types, arousal levels seem to not differ between checkers and non-checkers.

![Figure 3.2](image)

*Figure 3.2 mean arousal ratings for positive, neutral, negative, and checking-related selected pictures displayed separately for checkers and non-checkers. Higher scores represent less arousal. Error bars represent standard error measures. Neutral ratings have a score close to 5.*

The ANOVA showed a significant main effect of Picture Type, $F(1.68,33.64) = 14.46$, $p < .001$, but no significant interaction between the factors Picture Type and Group, $F(1.68,33.64) = 0.44$, $p = .61$. Paired sample post-hoc t-tests comparing picture types showed that arousal was enhanced for positive pictures ($4.37 \pm 0.32$) compared to neutral pictures ($5.47 \pm 0.16$, $t(21) = 3.96$, $p < .001$) and for negative pictures ($3.83 \pm 0.22$) compared to neutral pictures; $t(21) = 9.02$, $p < .001$. Emotional negative and positive pictures elicited similar arousal levels, $t(21) = 0.15$, $p = .13$. Checking-related pictures ($4.59 \pm 0.18$) elicited less arousal than negative pictures ($t(21) = 5.78$, $p < .001$), more arousal than neutral pictures ($t(21) = 6.62$, $p < .001$), and similar arousal levels compared to positive pictures ($t(21) = 0.66$, $p = .51$).
3.3.1.3. Picture Comfort Rating Analysis

Figure 3.3 shows comfort ratings for all picture types. As expected, positive pictures were perceived as more comfortable compared to negative pictures with intermediate scores for neutral pictures. Comfort ratings of checking-related pictures seemed to be more negative than for neutral pictures but not as negative as for negative pictures Group comfort ratings did not seem to differ between positive and neutral types. In contrast, checkers seemed to have more negative ratings for checking-related and negative pictures compared to non-checkers.

![Figure 3.3 mean comfort ratings for positive, neutral, negative, and checking-related selected pictures displayed separately for checkers and non-checkers. Lower scores represent more comfort and higher scores mean less comfort. Neutral ratings have a score close to 5.Error bars represent standard error measures.](image)

The ANOVA showed a significant main effect of Picture Type, $F(3, 60) = 309.59, p < .001$, and a significant interaction between the factors Picture Type and Group, $F(3, 60) = 6.80, p < .001$. Post-hoc paired sample t-tests showed that positive pictures ($2.36 \pm 0.15$) were perceived as more comfortable than neutral pictures ($4.28 \pm 0.15$, $t(21) = 10.61, p < .001$), neutral pictures were perceived as more comfortable than negative pictures ($7.23 \pm 0.16$; $t(21) = 18.79, p < .001$); and positive pictures were more comfortable than negative pictures, $t(21) = 20.28, p < .001$. Checking-related pictures ($5.72 \pm 0.18$) were rated as more comfortable than negative pictures, $t(21) = 10.29, p < .001$, but as less comfortable compared to neutral pictures ($t(21) = 8.79, p < .001$) and positive pictures ($t(21) = 15.02, p < .001$). Independent t-tests comparing picture types between groups did not reveal any group
difference for the positive and neutral picture types. However, there was group differences for negative pictures ((Mean (Checker): 7.66 ± 0.18; Mean (Non-checker): 6.8 ± 0.22; \( t(20) = 3.07, p_b = .01, r = .56 \)) and for checking-related pictures ((Mean (Checker): 6.35 ± 0.18; Mean (Non-checker): 5.09 ± 0.16; \( t(20) = 5.13, p_b = .002, r = .75 \)). For both picture types, checkers were more uncomfortable than non-checkers.

### 3.3.2. Analysis of all picture validation tasks from study 1 to 6

#### 3.3.2.1. Picture Valence Rating Analysis

Figure 3.4 shows the valence ratings for the picture types, considering all picture validation task data together. Positive pictures presented a lower valence compared to negative pictures, with intermediate scores for neutral pictures. Group valence ratings did not differ between these picture types. In the other hand, checkers seemed to have more negative ratings for checking-related pictures compared to non-checkers. Valence ratings of checking-related pictures seemed to be more negative than for neutral pictures but not as negative as for negative pictures.

![Figure 3.4 Mean valence ratings for positive, neutral, negative, and checking-related pictures considering all picture validation task studies from studies 1 to 6 together. Lower scores represent more positive valence and higher scores mean more negative valence. Neutral ratings have a score close to 5. Error bars represent standard error measures.](image)

As expected, the ANOVA showed a significant main effect of Picture Type, \( F(3, 15) = 1283.47, p < .001 \), and a significant interaction between the factors Picture Type and
Group, $F(3, 60) = 5.20, p < .003$. The significant main effect of Picture type was not investigated here, as previous analysis in studies 1 to 6 showed that positive pictures were always rated more positively than neutral pictures and neutral pictures were always rated more positively compared to negative pictures. Besides, negative pictures were always rated as more negative than positive pictures. Finally, checking-related pictures were always rated more positively than negative pictures and more negatively than neutral pictures and positive pictures. Independent t-tests comparing picture types between groups did not reveal any group difference for the positive, neutral, and negative picture types. However, there was a significant difference for the checking-related pictures between checkers and non-checkers ((Mean (Checker): 6.03 ± 0.12; Mean (Non-checker): 5.43 ± 0.07; $t(109.65) = 4.32, p_b = .001, r = .38$).

3.3.2.2. Picture Arousal Rating Analysis

Figure 3.5 shows that arousal ratings varied between picture types. As expected, it seemed that positive pictures generated less arousal compared to negative pictures but similar arousal compared to checking-related pictures. In addition, the arousal levels for positive pictures were higher than for neutral pictures. For all picture types, arousal levels seem to not differ between checkers and non-checkers.

![Arousal ratings](image)

Figure 3.5 mean arousal ratings for positive, neutral, negative, and checking-related considering all picture validation task studies from studies 1 to 6 together displayed separately for checkers and non-checkers. Higher scores represent less arousal. Error bars represent standard error measures. Neutral ratings have a score close to 5.
The ANOVA showed a significant main effect of Picture Type, $F(1.99, 254.84) = 69.44, p < .001$, but no significant interaction between the factors Picture Type and Group, $F(9.04,254.84) = 0.67, p = .74$. The main effect of Picture Type was not furtherly investigated, as it was already been shown that arousal was enhanced for positive pictures compared to neutral pictures and for negative pictures compared to neutral pictures in all previous analysis. Additionally, it has been shown that emotional negative and positive pictures elicited similar arousal levels; checking-related pictures elicited less arousal than negative pictures, more arousal than neutral pictures and similar arousal levels compared to positive pictures.

3.3.2.3. Picture Comfort Rating Analysis

Figure 3.6 shows comfort ratings for all picture types. Here, positive pictures were perceived as more comfortable compared to negative pictures with intermediate scores for neutral pictures. Comfort ratings of checking-related pictures seemed to be more negative than for neutral pictures but not as negative as for negative pictures Group comfort ratings did not seem to differ between positive and neutral types. In contrast, checkers seemed to have more negative ratings for checking-related compared to non-checkers.

![Figure 3.6](image)

*Figure 3.6 mean comfort ratings for positive, neutral, negative, and checking-related considering all picture validation task studies from studies 1 to 6 together displayed separately for checkers and non-checkers. Lower scores represent more comfort and higher scores mean less comfort. Neutral ratings have a score close to 5. Error bars represent standard error measures.
The ANOVA showed a significant main effect of Picture Type, \( F(2.78, 356.67) = 970.67, p < .001 \), and a significant interaction between the factors Picture Type and Group, \( F(2.78, 356.67) = 6.87, p < .001 \). The main effect of Picture Type was not furtherly investigated, as it was repeatedly been shown that positive pictures were perceived as more comfortable than neutral pictures, neutral pictures were perceived as more comfortable than negative pictures and positive pictures were more comfortable than negative pictures. Checking-related pictures were rated as more comfortable than negative pictures but as less comfortable compared to neutral pictures and positive pictures.

Independent t-tests comparing picture types between groups did not reveal any group difference for the positive, neutral and negative picture types. However, there was group differences for checking-related pictures \((\text{Mean (Checker): } 6.14 \pm 0.12; \text{Mean (Non-checker): } 5.22 \pm 0.11; t(138) = 5.59, p_b = .001, r = .43)\). For both picture types, checkers were more uncomfortable than non-checkers.

3.3.1.4. Summary and discussion of the selected picture analysis

This validation study aimed to evaluate the valence, arousal and comfort levels of the checking-related pictures in relation to the pictures selected from IAPS. Additionally, it aimed to investigate picture rating differences between checkers and non-checkers and it was expected that (1) checkers would rate checking-related pictures as more negative than non-checkers and (2) checkers would rate checking-related pictures as more negative than general negative pictures. Based on Simon et al. (2010) study, it was also expected that positive and neutral pictures would be similarly rated by checkers and non-checkers.

Note, only a selection of 100 pictures \((20 \text{ positive}, 20 \text{ negative}, 20 \text{ checking-related}, \text{ and } 40 \text{ neutral pictures})\) was needed for the binary decision making study. In each study from studies 2 to 6, the picture validation task was applied before the binary decision making task, unless at study 1, where only the picture validation task was present. Considering the huge amount of participants selected, the picture validation task data from all studies was pooled considering group and experiment and between-participant variables.

The picture selection was based on the valence ratings because valence is the most commonly used criterion to assess emotional directionality \((\text{e.g. Lang et al., 2008})\) and because the comfort scale is newly introduced in this study. The exclusion criteria were described in the methods section and in Table 3.3. The aim of this picture selection was to reduce group
differences for positive, neutral, and negative pictures, and to enhance group differences for checking-related pictures. Picture selection was conducted to (1) reduce the number of pictures for the binary decision making task, (2) to increase checker and non-checker valence rating differences for checking-related pictures while reducing group differences for positive, neutral, and negative pictures, and (3) to exclude neutral pictures with odd ratings, e.g. initially neutral pictures that were rated as positive or negative.

Firstly, considering only study 1, results have shown that valence ratings for positive, neutral and negative pictures were in accordance to their initial categorisation and there were no group differences for these categories. In addition, checking-related pictures were rated as more negative than neutral and positive pictures but were less negative compared to negative pictures for both groups. More importantly, checking-related pictures were rated as more negative by checkers compared to non-checkers. This is in accordance to findings by Simon et al. (2010) who showed that medication-free OCD patients were especially sensitive to emotional pictures associated with their obsessive compulsive subtype. Please note, these patients also had higher anxiety and depression comorbidities. We could confirm similar valence rating findings even for subclinical checkers who did not differ in their depression and anxiety ratings compared to non-checkers.

Secondly, the comfort scale was introduced as a measure because uncomfortable states could reflect the presence of obsessions (Veale et al., 1993; Veale, 2007 for a review). Interestingly, the ratings on the comfort scale were very similar to the valence ratings but they were not the same in study 1. More specifically, group differences were found for negative and checking-related pictures with more negative comfort ratings being given for both picture types by checkers compared to non-checkers. Hence, this might reflect a more general emotional response and not only a symptom-specific response. This enhanced emotional response to negative pictures cannot be explained by enhanced anxiety or depression levels because the groups did not significantly differ for anxiety and depression. Therefore, it is likely to reflect an enhanced emotional reactivity to general aversive affect materials in subclinical checkers. Regarding the arousal scale, there was no difference between groups for any of the picture types.

Considering the analysis between all studies, results were very similar for the arousal and valence scale, where checkers presented a higher and more negative score for valence towards the checker-related pictures compared with the non-checker group. For the arousal scale, no significant effects were found, as in study 1 but for the comfort scale, analysis has
shown that checkers and non-checkers similarly rated the negative pictures, while checkers rated the checker-related pictures as more uncomfortable than non-checkers, confirming results from study 1 only.

The arousal results contradict Simon et al. (2010) results, where symptom-related pictures generated more arousal in the OCD group than in their control group. Again, it is possible that our findings were different because our sample was subclinical and because participants in our checker and non-checker group had similar low levels of depression and anxiety. Enhanced anxiety might increase arousal ratings for all emotional pictures and for symptom-related emotional pictures specifically, as anxiety is usually related with a hyperarousal to emotional content (Mennin, Heimberg, Turk & Fresco, 2005).

In summary, group differences between subclinical checkers and non-checkers were found for the valence ratings of checking-related pictures in study 1 and for all the studies together. For the comfort ratings, group differences between subclinical checkers and non-checkers were found for the negative and checking-related pictures in study 1 and only for the checker-related pictures considering all studies together. Actually, this last analysis confirms that subclinical checkers seem to differently rate checker-related pictures compared to non-checkers, considering valence and comfort but it is not clear if they differently rate negative pictures when compared with non-checkers, as this result was only the case for study 1. The results from the valence scale were similar to the findings of the previous literature (Simon et al., 2010) although we employed subclinical checkers and not OCD patients. These group differences cannot be explained by anxiety and depression comorbidities.
Chapter 4: Pilot experiments

4.1. Introduction

Previous literature have shown that OCD presents a specific deficit in reward-based learning tasks, also known as decision making under ambiguity (Starcke et al., 2010; Kim et al., 2015; Zhang et al., 2015). Tasks as the Iowa Gambling Task (IGT) (Starcke et al., 2010; Kim et al., 2015) and the Frank task (Gründler et al., 2011) were used to investigate these deficits in OCD. However, results are contradictory as studies employing the IGT found OCD participants chose the disadvantageous options more often than controls (Starcke et al., 2010; Kim et al., 2015), while one study employing the Frank task has not found any deficit, i.e. OCD participants presented a similar performance to controls (Gründler et al., 2011).

It is possible that these contradictory findings are related with the tasks employed in these studies, as they simultaneously manipulate different factors that can affect decision making under ambiguity, such as feedback direction (positive feedback vs. negative feedback) and feedback magnitude (more vs. less positive feedback), for example. In this way, it is not possible to know which of these factors are in deficit in OCD.

The binary decision making task, disguised as a gambling task, was used to disentangle these reward factors and to try to investigate which of them is deficient in subclinical checkers. This task was designed in a way that we could study different feedback profiles and investigate different factors, such as the manipulation of feedback direction (positive-neutral vs. negative-neutral experiments), the manipulation of feedback magnitude (positive-neutral vs. positive-positive experiment or negative-neutral vs negative-negative experiment), and the alteration of the feedback magnitude by symptom-related information (negative-neutral vs. symptom-related experiment). In this way, five experiments (studies 2 to 6) with distinct feedback profiles were specifically designed for this PhD thesis. Each experiment had a particular emotional context.

Another methodological short-coming of the previous studies, which examined reward-based decision making in probabilistic environments in OCD patients, is that they restricted their analysis to accuracy measures only. However, learning rates can be different between two populations, even if the final result of accuracy is the same at the end. Additionally, it is common that participants use previous feedback to make actual decisions in RBLTs (Wilder et al., 2009), sometimes considering the only use of the immediate previous feedback (first order effect) or sometimes considering the use of more previous
feedbacks (second order effects) (Wilder et al., 2009). Therefore, it was decided to add new analyse measures to the investigation of RBL in subclinical checkers in this current PhD thesis, such as the analysis of learning curves (accuracy per sub-blocks), of the use of immediate previous trial information for the current decision (win shift, lose shift and cross-correlation values at lag 1), and second order effects (cross-correlations at lags 2-5).

Finally, the investigation of pattern sequences is another new addition to this PhD thesis and the wider research in OCD deficits in decision making and RBL. Therefore, we included pattern sequence blocks to our experiments. More specifically, each particular experiment was divided in three blocks, each one presenting a distinct probabilistic environment. In block 1, there were no sequential effects (no pattern) but a higher probability of winning in one of the target options. In block 2, a pattern sequence was presented, e.g. LLRR, but additional noise was added to make the detection of the pattern more difficult. Both target options had an equal probability of winning and performance could only be improved by implicitly learning the pattern in the event sequence. In block 3, there was a pattern and differential target option probabilities which were the same as in block 1. This block was designed to examine if OCD would be biased towards the pattern strategy or target option strategy usage.

Note that these tasks were not previously tested in any other study. Therefore, it was important to examine different sets of probabilities for each block to guarantee that participants would be able to learn each block in stages, so that learning progress can be examined. This pilot study had three aims regarding this issue: (1) to investigate distinct sets of sequences with distinct probabilities to examine whether participants would be able to learn in each block (2) to make sure that the learning process in each block happened in stages, so the event sequences should neither be too easy nor too hard to learn, (3) to test if new analysis techniques employed in later chapters could be used to measure accuracy changes over time (sub-block analysis), and pattern learning in the sequence (auto- and cross-correlations).

4.2. Methods

4.2.1. Overview of the study design

The binary decision making task employed here was already described at Chapter 2. Block 1 sequence was an order 0 Markov chain, block 2 was an order 2 and block 3 was
another order 2 Markov chain. Optimal strategies for each block did not vary between pilots and are described in Chapter 2. In fact, pilots varied in their probabilities between options and conditional probabilities. The manipulation of probabilities between options influenced the maximum accuracy values of a trial by trial strategy (no sequential effects). The manipulation of conditional probabilities influenced the maximum accuracy values of a pattern strategy. Experiments also varied in the number of participants and number of trials. Table 4.1 shows these information from pilots 1-5. Number of trials were decreased because participants were too tired in pilot 1 after the completion of the experiment. Maximum accuracy values were either increased because participants were not able to learn the task, or decreased because the pattern was too easy to learn, as participants reached the learning plateau at the first sub-block of 32 trials (higher maximum accuracies). Details about this manipulation are explained in each pilot subsection.

In pilot 1, there were two possible experiments participants could be randomly assigned: negative-negative experiment and positive-positive experiment. In the positive-positive experiment, participants received money and saw a positive picture after finding the target, while an absence of money and of a positive picture was followed after an error. In the negative-negative experiment, participants lost money and saw a negative picture after an error and an absence of penalty and of a negative picture was followed after participants found the target (see Chapter 2 for more details).

In pilots 2-5, only the positive-positive experiment was employed (Table 4.1). Depending on the experiment, participants either received a Participant information sheet (PIS) regarding the negative experiment or a PIS regarding the positive experiment (Appendix 2.4 for an example of PIS). After this, participants should read and sign the Consent form (Appendix 2.5). Participants received 10 pounds for participation plus the money they won at the decision making task. The total lab-based study took around 1 hour and a half and all the tasks were executed at one booth at the Department of Psychology at University of Surrey.
Table 4.1

**Characteristics of experiments for each pilot, containing number of participants, number of trials, values of target probability between options and maximum accuracy for the two possible strategies, investigated in this study.**

<table>
<thead>
<tr>
<th>Pilot number</th>
<th>Positive experiment</th>
<th>Negative experiment</th>
<th>Number of trials / block</th>
<th>Block 1</th>
<th>Block 2</th>
<th>Block 3</th>
<th>Trial by trial strategy</th>
<th>Pattern strategy</th>
<th>Trial by trial strategy</th>
<th>Pattern strategy</th>
<th>Trial by trial strategy</th>
<th>Pattern strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>5</td>
<td>200</td>
<td>0.66</td>
<td>0.5</td>
<td>0.66</td>
<td>0.66</td>
<td>0.5</td>
<td>0.5</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0</td>
<td>160</td>
<td>0.66</td>
<td>0.5</td>
<td>0.66</td>
<td>0.66</td>
<td>0.5</td>
<td>0.5</td>
<td>0.94</td>
<td>0.66</td>
<td>0.96</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>0</td>
<td>160</td>
<td>0.66</td>
<td>0.5</td>
<td>0.66</td>
<td>0.66</td>
<td>0.5</td>
<td>0.5</td>
<td>0.78</td>
<td>0.66</td>
<td>0.79</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>0</td>
<td>160</td>
<td>0.72</td>
<td>0.5</td>
<td>0.72</td>
<td>0.72</td>
<td>0.5</td>
<td>0.5</td>
<td>0.84</td>
<td>0.72</td>
<td>0.85</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>0</td>
<td>160</td>
<td>0.72</td>
<td>0.5</td>
<td>0.72</td>
<td>0.72</td>
<td>0.5</td>
<td>0.5</td>
<td>0.82</td>
<td>0.72</td>
<td>0.82</td>
</tr>
</tbody>
</table>
4.2.2. Participant selection

Participants were recruited among graduate and undergraduate students at the Department of Psychology of the University of Surrey. From pilots 1-4 there was no employment of the online questionnaire to screen participants, as, at the time the online questionnaire was being developed. Participants were only screened for comorbidities and checking symptoms in pilot 5. Ten University of Surrey undergraduate students filled out the online questionnaire. 7 participants were excluded, based on the online questionnaire. 3 participants participated in the binary decision making task. From the excluded participants, three participants presented depression, one did not want to take part in the binary decision making task, one gave up and two of them were neither checkers nor non-checkers, as their OCI-R scores for the checker sub-scale were between 2 and 5. At the end, from the three participants who executed the binary decision making task, two were non-checkers and one was a checker. Table 4.2 shows DASS and OCI-R scores for the participants who performed the binary decision making task in pilot 5

<table>
<thead>
<tr>
<th>Participants</th>
<th>DASS score</th>
<th>OCI-R score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Depression</td>
</tr>
<tr>
<td>P1 (checker)</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>P2 (non-checker)</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>P3 (non-checker)</td>
<td>22</td>
<td>8</td>
</tr>
</tbody>
</table>

4.2.3. Laboratory-based study

4.2.3.1. Data analysis for the laboratory-based study

Analysis used here were described in Chapter 2. Accuracy, sub-block accuracy and comparison of auto and cross-correlations were employed here for each pilot. Accuracy and sub-block accuracy were employed to investigate if participants were able to learn the task. Accuracies closer to 0.5, for example, indicated that participants were not able to learn the task, as this indicates random choices. Additionally, if accuracies did not increase between sub-blocks at the learning curve analysis, this would suggest that participants were not able to learn the optimal strategy in the task. In fact, as the previous analysis do not indicate if the participant
was able to learn the pattern in block 3 (as two strategies were possible), a cross-correlation analysis from lags 1-5 was also conducted. The cross-correlation analysis was not conducted for block 1, as the main focus of this chapter was to investigate if the conditional probabilities postulated for blocks 2 and 3 were creating a pattern with a specific value of noise (disruptions) that allowed learning. For the others chapters, however, this analysis was performed to investigate distinct decision making strategies in block 1.

It was expected that the experimental sequence presented values of correlation with a peak at lags 2 and 4 for block 2 and a peak at lag 3 for block 3. If participants learned the pattern in blocks 2 and 3, peaks of correlation were expected at these lags for the respective blocks. As described in the Methods in Chapter 2, the closer the participant cross-correlation values of the correlation values of the original sequence, the more the participants followed the pattern structure. If there were no peaks of cross-correlation at lags 2 and 4 and lag 3 for blocks 2 and 3 or if peaks were too distant from the correlation values of the original sequence, this indicated that participants did not learn or did not want to follow the pattern structure.

After calculating these values for cross-correlation and sequence auto-correlation, the presence of correlation peaks were investigated. Sequences where the participant cross-correlation values were too close to the experimental sequence were discarded. Sequences where the participant was too far from the experimental sequence were also discarded.

4.3 Piloting sequences (Results)

4.3.1 Pilot 1

4.3.1.1. Results for pilot 1

4.3.1.1.1. Accuracy

Independent t-tests between experiments showed there was no statistically significant difference between experiments for blocks 1, 2 and 3. Figure 4.1a shows that accuracy values were different from 0.5 in the positive-positive experiment (Mean: 0.54 ± 0.02) and negative-negative experiment (Mean: 0.58 ± 0.04) for block 1. Figure 4.1b shows that accuracy values were closer to 0.5 for the positive-positive experiment (Mean: 0.49 ± 0.03) and negative-negative experiment (Mean: 0.50 ± 0.02) for block 2. Figure 4.1c shows accuracy was different from 0.5 in the positive-positive experiment (Mean: 0.58 ± 0.02) and negative-negative experiment (Mean: 0.61 ± 0.02) for block 3.
Figure 4.1 Mean accuracy for the positive and negative experiments in pilot 1. Maximum accuracy is represented in black. A) Mean accuracy for block 1 with maximum accuracy of 0.66; B) Mean accuracy for block 2 with maximum accuracy of 0.66; C) Mean accuracy for block 3 with maximum accuracy of 0.66
In summary, as accuracy was close to 0.5 in block 2, this indicated participants were not able to learn the optimal strategy in this block. However, for blocks 1 and 3, more analysis, as the sub-block accuracy, are needed to confirm if participants were able to learn the optimal strategy in the task.

4.3.1.1.2. Sub-block accuracy (learning effect)

A two-way 4 x 2 mixed ANOVA was performed with the within-participant factor sub-block (sub-blocks 1-5) and the between-participant factor experiment (positive-positive vs. negative-negative) for blocks 1, 2 and 3. Post-hoc independent t-tests were conducted and Bonferroni corrections applied when necessary.

The ANOVA revealed that there was no significant main effect of experiment and interactions for blocks 1, 2 and 3. Especially, there was no significant main effect of sub-block for each block, so accuracy values did not variate between sub-blocks (Figure 4.2a for block 1; 4.2b for block 2 and 4.2c for block 3). In this way, results indicated that participants were not able to increase accuracy across sub-blocks, which could suggest participants were not able to learn the optimal strategy for blocks 1, 2 and 3.
Figure 4.2 Mean accuracy in each sub-block of 40 trials for the positive-positive and negative-negative experiments in pilot 1. A) Mean accuracy for block 1; B) Mean accuracy for block 2; C) Mean accuracy for block 3
4.3.1.1.3. Comparison of auto-and cross-correlations

A three-way 5 x 2 x 2 mixed ANOVA on the auto- and cross-correlation difference scores, considering lags as a within-participant factor and experiments as a between-participant factor was conducted for each block. There was no main effect of lags, experiment or interaction for the three blocks (Figures 4.3a and b).

It was expected that participants would present a higher value of negative cross-correlation at lag 2 and positive cross-correlation at lag 4 for block 2. Additionally, it was expected a higher value at lag 3 (pattern strategy) for block 3. To facilitate visualization of how cross-correlation values should be in relation with the experimental sequence, the averaged auto-correlation values of the experimental sequence were plotted together with the participant’s cross-correlation values Figures 4.3a and b show that participants were not able to learn block’s 2 and 3 pattern strategy, as they did not seem to be cross-correlating their responses with lags 2 and 4 in block 2 or lag 3 in block 3.
Figure 4.3 Cross-correlation probabilities from lags 1 to 5 between the positive-positive and negative-negative experiments in pilot 1. The averaged auto-correlation values of the experimental sequence are shown in black. A) Block 2 cross-correlation values; B) Block 3 cross-correlation values

4.3.1.2. Discussion of pilot 1

Participants

Participants reported that the experiment was too long. In this way, the number of trials was decreased in pilot 2.
Differences between experiments

There were no significant differences between the negative-negative and the positive-positive experiments for blocks 1, 2 and 3, regarding all analysis employed. It is possible that the small sample size and the fact that non-checkers and checkers were not screened by questionnaires were responsible for the absence of non-significant findings between experiments.

Experimental sequences

Block 1

Accuracy values indicated that participants were learning the task, as accuracy was different from 0.5, suggesting random choices. However, sub-block accuracy analysis did not revealed any difference between accuracies across distinct sub-blocks. It is possible participants learnt the optimal strategy at the first sub-block and continued to employ it until the end of the block. This could explain why accuracy values did not vary between sub-blocks.

Block 2

For block 2, results revealed that participants were not able to learn the optimal strategy. Firstly, accuracy levels were about 0.5, suggesting random choices. Moreover, there was no evolution in accuracy across sub-blocks (learning effect analysis), other indication participants were not learning anything. Finally, cross-correlation analysis showed that values for all the previous trials were close to zero for both experiments. These results suggest that participants were not able to learn the pattern in the task. In this way, it seems the conditional probabilities used in this block were not able to generate a learnable pattern here. Thus, a new set of sequences were developed for block 2 to be tested at pilot 2.

Block 3

Accuracy and sub-block accuracy revealed that participants were able to learn an optimal strategy from block 3. However, the cross-correlation analysis indicated that participants were not able to learn the pattern in this sequence, as the cross-correlation value at lag 3 was close to zero. As the accuracy values were above 0.5, this indicates participants were opting for block 1 strategy in block 3. It is possible the pattern in block 3 was too hard to learn, as was the case of block 2. In this way, it seems participants opted for the easier strategy in
block 3. At the end, the set of sequences for block 3 was also modified and tested at the subsequent pilots.

In summary, as participants were too tired at the end of the task, the number of trials was decreased to 160 per block. Additionally, as results in pilot 1 showed that participants were unable to learn the patterns at blocks 2 and 3, it was decided to increase the values of maximum accuracy for these blocks, regarding the pattern strategy (table 4.1). The probability between options for block 1 was not modified, as it seemed participants were able to learn this optimal strategy, especially in block 3.

4.3.2. Pilot 2

Methods were similar to pilot 1. Only conditional probabilities for blocks 2 and 3 were changed, changing maximum accuracy values (table 4.1). The noise in these sequences was decreased, so there were less disruptions in the patterns. Only two participants were selected (Table 4.1), so results will only show individual scores for each participant.

4.3.2.1. Results of pilot 2

4.3.2.1.1. Accuracy

Figure 4.4a revealed that accuracy values were similar to 0.5 for participant 1 but different from 0.5 for participant 2 in block 1, indicating that one participant was able to learn this optimal strategy. Figure 4.4b showed that accuracy values were closer to maximum accuracy values for both participants, indicating they were able to learn the optimal strategy in block 2. Finally, figure 4.4c revealed that accuracy for both participants was similar to maximum accuracy values of block 3. Here, it is not possible to know which optimal strategy participants were following in this block.
Figure 4.4 Accuracy for two participants in pilot 2. Maximum accuracy is represented in black. A) Accuracy for block 1 with maximum accuracy of 0.66; B) Accuracy for block 2 with maximum accuracy of 0.94; C) Accuracy for block 3 with maximum accuracy of 0.96
4.3.2.1.2. **Sub-block accuracy**

Figure 4.5 revealed sub-block accuracy values for two participants in pilot 2. Figure 4.5a indicated both participants were able to learn the task in block 1, as accuracy increases across sub-blocks. Figures 4.5b and c indicated both participants were able to learn the task in blocks 2 and 3 at the first sub-block.

*Figure 4.5. Accuracy in each sub-block of 32 trials for two participants in pilot 2. A) Accuracy for block 1; B) Accuracy for block 2; C) Accuracy for block 3*
4.3.2.1.3. Comparison of auto and cross-correlation

Figures 4.6a and 4.6b showed that participants were able to learn the optimal strategy in blocks 2 and 3. This supports sub-block accuracy findings.

*Figure 4.6* Cross-correlation probabilities from lags 1 to 5 for two participants in pilot 2. The averaged auto-correlation values of the experimental sequence are shown in black. A) Block 2 cross-correlation values; B) Block 3 cross-correlation values
4.3.2.2. Discussion of pilot 2

For Block 1, participants had a similar performance when compared with pilot 1, what was expected, as the probabilities hadn’t changed. For blocks 2 and 3, however, sub-block accuracy results showed that participants were able to reach maximum accuracy after the first 32 trials (first sub-block). After analysing the cross-correlation data for blocks 2 and 3, it was possible to notice that participants were completely able to cross-correlate their answers with the second and fourth previous trial in block 2 and with the third previous trial in block 3. This indicated participants were able to learn conditional probabilities in blocks 2 and 3 and consequently, follow the pattern in the sequence. So, contrary to pilot 1, participants were able to learn the optimal strategy very quickly. Interestingly, participants preferred to follow the pattern structure in block 3 than to follow block 1 strategy, as it happened in pilot 1.

One issue of these results regards the fact that the optimal strategy was so easy to learn that participants reached the maximum accuracy at the first sub-block (figure 4.5) in blocks 2 and 3. As the sequences were designed with the aim to generate a learning curve with a mild learning rate, blocks 2 and 3 sequences were not ideal, as they were too easy to learn. In this case, maximum accuracies and consequently the probabilities of blocks 2 and 3 were again modified. Pilot 3 shows results for the new maximum accuracy values.

4.3.3. Pilot 3

The same method for the generation of experimental sequences was used in this pilot. Only conditional probabilities for blocks 2 and 3 were modified, changing maximum accuracy values (table 4.1). The noise in these sequences was increased, so there were more disruptions in the patterns. As in pilot 2, only two participants were selected (Table 4.1), so results only show individual scores for each participant.

4.3.3.1. Results of pilot 3

4.3.3.1.1. Accuracy

Figure 4.7a revealed that accuracy values were similar to 0.5 for participant 1 and participant 2 in block 1, indicating that no participant was able to learn its optimal strategy. Figure 4.7b showed that accuracy values were similar to 0.5 for one participant and above 0.5 for the other participant in block 2, so it seems one participant was able to learn the task in
block 2. Finally, figure 4.7c revealed that accuracy for both participants was similar to 0.5 in block 3, suggesting that no participant was able to learn any of the two strategies in block 3.

Figure 4.7 Accuracy for two participants in pilot 3. Maximum accuracy is represented in black. A) Accuracy for block 1 with maximum accuracy of 0.66; B) Accuracy for block 2 with maximum accuracy of 0.79; C) Accuracy for block 3 with maximum accuracy of 0.78
4.3.3.1.2. Sub-block accuracy (learning effect)

Figure 4.8 revealed that one participant was able to learn the optimal strategy in the last sub-block but the other participant did not present progression in accuracy across sub-blocks (Figure 4.8a). In block 2, it seems both participants were able to learn the task but their accuracies decreased after the second sub-block (Figure 4.8b). Finally, in block 3, both participants were not able to learn any of the optimal strategies in the task (Figure 4.8c).

Figure 4.8 Accuracy in each sub-block of 32 trials for two participants in pilot 3. A) Accuracy for block 1; B) Accuracy for block 2; C) Accuracy for block 3
4.3.3.1.3. Comparison of auto and cross-correlation

Figure 4.9 supports sub-block accuracy results about learning in blocks 2 and 3, as both participants did not reach expected correlation values for lags 2 and 4 in block 2 and lag 3 in block 3 (figure 4.9a and 4.9b).

*Figure 4.9* Cross-correlation probabilities for lags 1 to 5 for two participants in pilot 3. The averaged auto-correlation values of the experimental sequence are shown in black. A) Block 2 cross-correlation values; B) Block 3 cross-correlation values
4.3.3.2. Discussion of pilot 3

Sub-block accuracy results revealed that maximum accuracy changes in blocks 2 and 3 were not enough to generate an environment where participants would be able to learn the optimal strategy at a mild learning rate. Additionally, cross-correlation results confirmed that participants were not able to correlate their responses with lags 2 and 4 in block 2 and lag 3 in block 3. Previously, a maximum accuracy of 0.66 was employed in pilot 1 but this value was not ideal; in pilot 2, the maximum accuracy was increased to 0.94 and 0.96 for blocks 2 and 3, respectively but the optimal strategy was too easy to learn. Finally, in pilot 3, maximum accuracies of 0.78 and 0.79 for blocks 2 and 3 presented similar results to pilot 1. In this way, the maximum accuracies of 0.78 and 0.79 for both blocks, respectively, were increased again. As results from pilot 2 had already shown that maximum accuracies of about 0.95 were not ideal, a set of sequences with maximum accuracy values between 0.78 and 0.94 for block 2 and 0.79 and 0.96 for block 3 was designed for pilot 4.

4.3.4. Pilot 4

The same method of generating experimental sequences was used in this pilot. Conditional probabilities for blocks 2 and 3 were changed, modifying maximum accuracy (table 4.1). The noise in these sequences was decreased, so there were less disruptions in the patterns. After changing conditional probabilities for block 3, it was not possible to keep the probability between options as 0.66 for this block. To keep probabilities between options similar between blocks 1 and 3, the probability between options for blocks 1 and 3 was increased to 0.72. Four participants were selected for this pilot (Table 4.1), so results will only show the average and standard errors without any inferential statistic test.

4.3.4.1. Results of pilot 4

4.3.4.1.1. Accuracy

Figure 4.10a revealed that accuracy values were above 0.5 in block 1, indicating that participants were not choosing options randomly. Figure 4.10b and c showed that accuracy values were also above 0.5 for blocks 2 and 3.
Figure 4.10 Accuracy in pilot 4. Maximum accuracy is represented in black. A) Accuracy for block 1 with maximum accuracy of 0.72; B) Accuracy for block 2 with maximum accuracy of 0.84; C) Accuracy for block 3 with maximum accuracy of 0.85
4.3.4.1.2. Sub-block accuracy (learning effect)

Figures 4.11a, 4.11b and 4.11c revealed there was a learning progression across sub-blocks for blocks 1, 2 and 3. However, it seems the learning progression for block 3 was faster than aimed.

Figure 4.11 Mean accuracy in each sub-block of 32 trials in pilot 4. A) Mean accuracy for block 1; B) Mean accuracy for block 2; C) Mean accuracy for block 3
4.3.4.1.3. Comparison of auto- and cross-correlation

Cross-correlation analysis indicated that participants were able to learn blocks 2 (Figure 4.12a) and 3 optimal strategies (Figure 4.12b), especially in block 3 where lag 3 cross-correlation value was almost the same as the sequence auto-correlation value.

Figure 4.12 Cross-correlation probabilities from lags 1 to 5 in pilot 4. The averaged auto-correlation values of the experimental sequence are shown in black. A) Block 2 cross-correlation values; B) Block 3 cross-correlation values
4.3.4.2. Discussion of pilot 4

Sub-block accuracy for block 1 revealed that participants were able to learn the task. Regarding blocks 2 and 3, participants were also able to learn the optimal strategy. Additionally, in block 3, participants chose to follow the pattern strategy.

Regardless of the fact participants were performing as expected in blocks 1, 2 and 3, participants were learning the optimal strategy too fast and efficiently for block 3. In this way, maximum accuracy values for block 3 should be decreased in order to slightly increase the noise in the pattern of this sequence. In fact, this study aimed to create an experimental design where maximum accuracies for the between option strategy (Blocks 1 and 3) and for the pattern strategies (blocks 2 and 3) were as similar as possible. In this way, as maximum accuracy for block 3 needed to be decreased, the maximum accuracy value for block 2 was also decreased, even though results from pilot 4 were ideal for block 2. In this way, a new set of sequences was designed where maximum accuracies for blocks 2 and 3 were decreased to 0.82. Block 1 maximum accuracy was kept, as the probability between options for block 3 continued to be 0.72. This new set of sequence was tested in pilot 5, only for blocks 2 and 3.

4.3.5. Pilot 5

Conditional probabilities for blocks 2 and 3 were changed, modifying maximum accuracy (table 4.1). The noise in these sequences was increased, so there were more disruptions in the patterns. Three participants were selected here (Table 4.1), so results will only show data for each participant.

4.3.5.1. Results of pilot 5

4.3.5.1.1. Accuracy

Figure 4.13a revealed that accuracy values were about 0.5 in block 2, indicating that participants were choosing options randomly. Figure 4.13b revealed that accuracy values were above 0.5 for block 3. In this case, participants effectively learnt the task in block 3.
Figure 4.13 Accuracy in pilot 5 for each participant. Maximum accuracy is represented in black. A) Accuracy for block 2 with maximum accuracy of 0.82; C) Accuracy for block 3 with maximum accuracy of 0.82

4.3.5.1.2. Sub-block accuracy (learning effect)

The sub-block accuracy analysis indicated that participants were not able to learn the optimal strategy in block 2 (Figure 4.14a). For block 3, participants were able to almost reach maximum accuracy (Figure 4.14b). Additionally, the maximum accuracy was only reached at the last sub-block, confirming that this maximum accuracy value is efficient to generate a sequence with mild learning rate in block 3.
Figure 4.14 Accuracy in each sub-block of 32 trials for two participants in pilot 5. A) Accuracy for block 2; C) Accuracy for block 3

4.3.5.1.3. Comparison of auto and cross-correlation

Confirming sub-block accuracy results, participants were able to learn the optimal strategy in block 2, as lags 2 and 4 values were close to zero (figure 4.15a). However, figure 4.15b showed that participants were able to learn the pattern strategy in block 3, as cross-correlation values for lag 3 indicated participants were using the pattern in the sequence as a strategy.
Figure 4.15 Cross-correlation probabilities from lags 1 to 5 for three participants in pilot 5. The averaged auto-correlation values of the experimental sequence are shown in black. A) Block 2 cross-correlation values; B) Block 3 cross-correlation values
4.3.5.2. Discussion of pilot 5

Analysis showed that the maximum accuracy of 0.82 was ideal to generate learning with a mild learning rate in block 3. However, participants were not able to learn the optimal strategy in block 2. The maximum accuracy of 0.82 used in pilot 5 was an intermediate value between 0.84 (pilot 4) and 0.79 (pilot 3). In pilot 4, the block 3 sequence with a maximum accuracy of 0.84 was too easy to learn. For block 2, however, participants were able to learn the pattern at a mild rate. For the maximum accuracy of 0.82 in pilot 5, participants were able to learn the sequence at a mild rate in block 3 but block 2 results were similar to the ones of pilot 3, indicating the maximum accuracy of 0.82 is not ideal for block 2.

To keep the maximum accuracies as equal as possible between blocks 2 and 3, there were two options here: to use the sequence set of pilot 4 but assuming that block 3 pattern would be too easy to learn or use pilot 5 sequences but risk the fact that participants would not be able to learn the optimal strategy in block 2 at a mild rate. The second strategy was selected. Indeed, it was considered that it was not desirable to have a sequence in block 3 that was too easy to learn. In this way, it would be better to have two sequences in blocks 2 and 3 that presented a challenge to learn than just one sequence alone. Pilot 4 results showed that participants were only challenged in block 2. However, pilot 5 maximum accuracy presented a challenge in both blocks. Clearly, there was a risk that participants would not be able to learn block 2 strategy at all. As there was no more time to continue the development of new pilots, as a prevention, if block 2 sequences continued to be a problem in studies with a larger sample, new sequence sets would be designed afterwards. However, as it will be seen in the subsequent chapters, participants were able to learn block 2 optimal strategy with the maximum accuracy of 0.82 with a mild learning rate. Hence, it is possible that learning in block 2 did not happen in pilot 5 because the sample size was too small.

4.4. General conclusion

After selecting the best pictures for the experiments in Chapter 3 and after choosing the best experimental sequence set (pilot 5) in this chapter, the investigation of the effect of feedback direction and feedback magnitude (including the presence of symptom-related emotional stimuli) was conducted. Chapter 5 shows the investigation of the effect of feedback direction (studies 2 and 3); Chapter 6 shows the investigation of the effects of feedback magnitude (studies 4 and 5) and Chapter 7 (study 6) shows the effect of the presence of symptom-related emotional stimuli in the reward-based learning of subclinical checkers.
Chapter 5: Reward based-learning in subclinical checkers: The influence of feedback direction

5.1. Introduction

As shown before, OCD population is characterized by many cognitive and emotional deficits that are related with the orbitofrontal and dorsolateral-striatal loop, as well as mesolimbic circuits (Menzies et al., 2008). Memory (Harkin & Kessler, 2011), set shifting and inhibition (Dittrich & Johansson, 2013), as well as decision making deficits (Starcke et al., 2010) are usually present in OCD. Specially, there were ambiguous results regarding this last topic. For decision making tasks under ambiguity, sometimes they report deficits and sometimes they do not (for example, see Starcke et al. (2010) versus Gründler et al. (2009), respectively). In decision making under risk, where target probabilities are known to the participant, OCD patients do not show any deficits (Starcke et al., 2010). In such tasks, OCD individuals usually presented the same frequency of correct decisions as the control group (see the Game of dice task (GDT) in Chapter 1.6.1).

However, as stated above, the picture is more complicated for decision making under ambiguity or reward-based learning tasks (RBLTs), where individuals need the use of feedback to learn about target option probabilities (Starcke et al., 2010; Kim et al., 2015; Zhang et al., 2015). Actually, in studies using RBLTs, OCD participants presented a lower accuracy than the control group (Starcke et al., 2010) or they usually presented a higher frequency of choosing the disadvantageous options in the Iowa Gambling task (IGT) (see Starcke et al., 2010; Kim et al., 2015). However, OCD participants presented the same frequency of advantageous choices in the Frank task (Gründler et al., 2011).

These ambiguous findings could be explained by several reasons, e.g. the use or absence of medication, the presence or absence of common comorbidities associated with OCD as depression and anxiety, the study of OCD in general without differentiation of subtypes, and the specific task used in these studies. In fact, these tasks often manipulate several factors at the same time and a systematic investigation of these factors could be beneficial for the understanding of OCD deficits in decision making under ambiguity / RBL. These factors are feedback direction (positive vs. negative vs. mixed feedback during learning), feedback magnitude (e.g. amount of positive vs. negative feedback), feedback gain (difference in feedback between options), or feedback delay (immediate vs. delayed reward). For example, clear deficits were found for OCD patients in the IGT task but they used mixed
positive and negative feedback of different feedback magnitudes and, therefore, also different feedback gains. Moreover, they had an immediate and a delayed feedback manipulation. The question is which of these factors is specifically linked to the OCD deficit?

Another methodological issue of the previous studies is that they restricted the analysis only to accuracy measures and not even learning over time (learning curves). Whilst accuracy is an important measurement of performance and an indirect measure of the optimal strategy use in a task, others measurements could prove very useful when examining disease related deficits in RBLTs. In fact, it is common that participants use the immediate previous trial feedback (first order effect) to make actual decisions in RBLTs (Wilder et al., 2009). Actually, reward-based learning models usually assert that reinforcement learning is based on prediction errors (Maia, 2009) that are a comparison between the unexpected immediate previous results of a decision with the hypothesized result that was based on an internal model generated by the participant (Maia, 2009). If prediction errors are so important for reward-based learning, new measurements of how participants employ previous trials to make actual decisions and, more important, how participants deal with unexpected previous feedback, could be very useful to examine deficits in RBL.

Unfortunately, the use of previous trial results by participants cannot be measured by accuracy, so new analysis investigating the use of previous trials could increase the understanding of the decision making deficits associated with OCD. For instance, it is possible to investigate how participants are influenced by immediate previous results measuring the probability of shifting across options after winning or losing in the previous trials. Additionally, it is possible to investigate how participants are affected by one or more previous trials, measuring the probability of correlation between one actual response to a previous response (cross-correlation).

Finally, another gap in the literature of RBL in OCD concerns the absence of the investigation of possible deficits of sequential effects in OCD. Previous studies on RBLTs have shown that participants use second, third and more previous feedbacks to make predictions, as they are supposedly searching for patterns (second order or sequential effects) (Vulkan, 2000; Unturbe & Corominas, 2007; Wilder et al., 2009). This effect might be also present in random sequence tasks (Vulkan, 2000; Unturbe & Corominas, 2007) and might affect decision making. However, it is better to investigate these sequential effects with event sequences that have hidden patterns to find out how people with OCD or subclinical OCD symptoms deal with these pattern sequences. Actually, this might be another potential
explanation for the already known decision making deficits of OCD but it might also reveal additional deficits or coping strategies.

In summary, considering the issues from the previous investigations, the present investigation decided to study a subclinical checker sample without medication. In addition, a RBLT, i.e. the binary decision making task, was used because feedback direction, feedback magnitude, etc. could be manipulated separately. Additionally, sequential effects were shown to influence decision making in random tasks (Unturbe & Corominas, 2007). Hence, our RBLT was designed to investigate how the checker group deals with sequences with patterns (presence of real sequential effects) in comparison with a sequence without patterns. Finally, new analysis were employed to investigate first and second order effects in OCD (second order effect) by measuring the probability of shifting after winning and losing after the immediately previous trial in a random sequence block, and auto- and cross-correlations with the five previous trials in all blocks (see Chapter 2 for methods details).

This Chapter aimed 1) to examine the effect of the manipulation of feedback direction on RBL in subclinical checkers and non-checkers, and 2) to investigate changes in strategy, in addition to accuracy, which are reflected in first- and second order sequential effects in the decision making of OCD. Random-sequence and pattern sequence blocks were used to investigate these effects.

For the first aim, two binary RBL experiments were designed. The positive-neutral experiment (study 2) aimed to examine the effect of positive feedback direction in OCD, where only positive monetary feedback was present after winnings and an absence of feedback after losses. The negative-neutral experiment (study 3) aimed to examine the effect of negative feedback direction in OCD, where an absence of money loss was present after winnings and a negative monetary feedback was present after losses. Note that Starcke et al. (2010) associative feedback task was a first attempt to segregate feedback direction in positive and negative feedback but this task did not present many trials to evaluate learning. They also offered either mixed feedback, as winning different amounts of feedback for each option in the positive task, or losing different amounts of feedback for each of the options (no absence of feedback). In contrast, our tasks included an absence of reward (positive-neutral as study 2) or an absence of penalty (negative-neutral as study 3) option (see Chapter 1.6.3 and Chapter 2 for more details). They also had more trials per block to explore differential learning effects over time.
For the second aim, to investigate sequential effects in OCD, each experiment was divided into three blocks, each one presenting a distinct probabilistic environment. In block 1, there were no sequential effects (no pattern) but a higher probability of winning in one of the target options. This block presented a probabilistic environment that is usually employed in the literature (for OCD studies, e.g. Starcke et al., 2010; Gründler et al., 2011; for general decision making studies, e.g. Vulkan, 2000). In block 2, a pattern sequence was presented, e.g. LLRR, but additional noise was added to make the detection of the pattern more difficult. Both target options had an equal probability of winning and performance could only be improved by implicitly learning the pattern in the event sequence. In block 3, there was a pattern and differential target option probabilities which were the same as in block 1. This block was designed to examine if OCD would be biased towards the pattern strategy or target option strategy usage.

Based on Starcke et al. (2010) results, it is possible that the checker group will present a similar deficit for both the positive and negative feedback direction. However, it is also possible that results will be different, as our task presented more trials per block and an absence of feedback option. In fact, it is possible that checkers will need more trials to learn the task correctly but at the end, they will present an accuracy that is similar to non-checkers, meaning the learning curve might be different between groups.

It is also possible that checkers will present more deficits in the negative-neutral (study 3) than in the positive-neutral experiment (study 2) because previous literature has shown that OCD is more sensitive to aversive stimuli than positive stimuli (Simon et al., 2010). The presence of penalties in the negative-neutral experiment might decrease the checker ability to make correct decisions in the negative-neutral experiment in comparison to the positive-neutral experiment, where there are no such penalties. If this is correct, it is possible that checkers will be more biased towards behaviours of risk-avoidance in the negative-neutral experiment, in a way to avoid receiving negative feedback after a decision. Because of that, checkers will be less biased towards exploration of new hypothesis in the three blocks, while non-checkers will present more exploration in this experiment. This effect might only appear in the negative-neutral experiment (study 3).

If checkers will be less biased towards exploration in the negative-neutral experiment, it is expected that in block 1, checkers will present a higher accuracy than non-checkers, because they will try to avoid shifting options after an error, so they will present more choices in the option with more targets than the non-checker group. In this case, they will
also present a lower probability of shifting after losing and higher auto- vs. cross-correlation
differences for lags 1 to 5 in block 1 than non-checkers, as they will repeat their choices in
the same option, increasing the chance of correlation with previous trials. There will not be
any difference in the number of shifts after a previous winning, as checkers are supposedly
more affected by penalties than rewards. Clearly, these results will be only the case, if
checkers will be able to learn that one option presents more feedback than the other.

In the case of blocks 2 and 3, if checkers are able to learn the pattern structure of the
event sequence, they will present a persistent behaviour of choosing the pattern, disregarding
pattern disruptions in the sequence to avoid penalties. This will be more the case for the
negative-neutral experiment (study 3), as penalties will decrease the bias towards exploration
of new hypothesis in this group. Hence, checkers will present a higher accuracy and more
cross-correlations with the previous trials associated with the specific pattern of these
sequences (i.e., second and fourth previous trials for block 2 and third previous trial for block
3 pattern). In block 3, it is possible that checkers will present a bias towards block 1 strategy
because it needs less exploration. If this is not the case, however, it is also possible that
results will be similar to the ones of block 2.

Finally, it is expected that measures of accuracy, lose-shift and cross-correlation will
be correlated with measures of checking habits, general obsessive-compulsive tendencies,
anxiety and depression in the negative-neutral experiment (study 3).

5.2. Methods

5.2.1. Overview of the study design

An overview about the general study design is presented in Figure 5.1. In this study,
students completed an online screening questionnaire to find potential subclinical checker
and non-checker participants for the laboratory-based part of the study. After participants
were selected and invited to the laboratory, they were randomly assigned to take part in either
the positive-neutral (study 2) or the negative-neutral experiment (study 3). In the laboratory-
based part of the study, all participants performed a picture validation task, a binary decision
making task (either positive-neutral or negative neutral task), and completed several
laboratory-based questionnaires. More details about the specifics of this study can be found
in Chapter 2 and in the Chapter section 5.2.
Figure 5.1 General study design for the positive-neutral (study 2) and negative neutral experiments (study 3)

5.2.2. Participant selection based on the online screening questionnaire

The Online screening questionnaire (Appendix 2.3) consisted of a demographic questionnaire asking about age, gender, videogame playing habits and gambling for money habits, past or present diagnosis of depression, anxiety, panic disorder and post-traumatic stress disorder, as well as the Obsessive Compulsive Inventory revised (OCI-R) and the Depression, Anxiety and Stress Scales (DASS) questionnaires. More details about the online screening questionnaire can be found in Chapter 2.

The online screening questionnaire was filled out by 340 students (see Table 5.1. for details about exclusions, etc.). However, 85 participants did not complete the online questionnaire, 1 participant did only answer with one response choice, and 11 participants filled out the questionnaire twice. The remaining completed 243 online questionnaires were analysed further. Of these 243 questionnaires, 125 participants were excluded on the basis of the study exclusion criteria and 118 participants were invited to the laboratory-based study. Of these invited participants, 49 participants completed the second part of the study.

During data analysis of the laboratory-based tasks, three additional participants were excluded, i.e. one participant only used one response key in the picture validation task in the second block and, therefore, was not doing the task as instructed. This participant was replaced. In addition, during the data analysis it was noticed that one participant in the positive-neutral non-checker group was actually a checker. This participant and the matching non-checker participant were excluded from the positive-neutral experiment (study 2), so groups could present the same number of participants. Hence, the final data analysis was
performed on 24 participants in the negative-neutral experiment (12 checkers and 12 non-checkers) and 22 participants in the positive-neutral experiment (11 checkers and 11 non-checkers).

Table 5.1

**Reasons and number of participants excluded and included in the study**

<table>
<thead>
<tr>
<th>Exclusion reason</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total participants</td>
<td>340</td>
</tr>
<tr>
<td>Incomplete questionnaires</td>
<td>85</td>
</tr>
<tr>
<td>Questionnaire completion incorrect (e.g. only one response type)</td>
<td>1</td>
</tr>
<tr>
<td>Questionnaire completed twice</td>
<td>11</td>
</tr>
<tr>
<td>Questionnaire completed</td>
<td>243</td>
</tr>
<tr>
<td>Age</td>
<td>1</td>
</tr>
<tr>
<td>Diagnosed anxiety disorder (controls)</td>
<td>8</td>
</tr>
<tr>
<td>Diagnosed depression</td>
<td>13</td>
</tr>
<tr>
<td>Diagnosed depression &amp; PTSD</td>
<td>1</td>
</tr>
<tr>
<td>Diagnosed depression &amp; anxiety</td>
<td>8</td>
</tr>
<tr>
<td>Diagnosed panic disorder</td>
<td>3</td>
</tr>
<tr>
<td>DASS score</td>
<td>40</td>
</tr>
<tr>
<td>Neither checker, nor non-checker</td>
<td>49</td>
</tr>
<tr>
<td>OCD symptoms but no checker subtype</td>
<td>2</td>
</tr>
<tr>
<td>Remaining participants</td>
<td>118</td>
</tr>
<tr>
<td>No response to invitation to lab-based study</td>
<td>69</td>
</tr>
<tr>
<td>Recorded for the lab-based study</td>
<td>49</td>
</tr>
<tr>
<td>Exclusion from the lab-based study</td>
<td>3</td>
</tr>
<tr>
<td><strong>Complete results</strong></td>
<td><strong>46</strong></td>
</tr>
</tbody>
</table>

5.2.3. **Participants in the laboratory-based study: Group sample characteristics**

For tests of normality, z-scores for skewness and kurtosis were calculated for age, and the DASS and OCI-R scores. DASS depression and DASS total score in the negative-neutral experiment were not normally distributed; the OCI-R total, OCI-R washing, OCI-R ordering, OCI-R obsessing scores in the negative-neutral experiment were not normally distributed;
and the OCI-R counting score in the positive-neutral experiment were not normally distributed.

For most variables, analysis were conducted by using between-participant ANOVAs with the factors Group (checker vs. non-checker) and experiment (positive-neutral experiment vs. negative-neutral experiment). Gender, videogame playing and gambling with money were the exceptions because these were categorical variables. Here, Chi-Square tests or Fisher’s exact tests were used to explore sample differences between groups or experiments. The findings from this analysis are presented in Table 5.2.

For both the OCI-R and the DASS scores, ANOVAs did reveal main effects of Groups (Table 5.2). Basically, checkers always had higher scores than non-checkers in all OCI-R and DASS subscales, independent of the experiments (Table 5.2). For age, however, non-checkers (22.47 ± 0.09) were slightly older than checkers (20.03 ± 0.05; main effect of Group: $F(1, 42) = 4.10, p =.04$). However, this difference is unlikely to affect performance, as previous studies showed young adults from 20-25 years old present similar decision making strategies (Victorino et al., unpublished).

There was no significant difference between experiments or groups for the gender, play videogame and play for money variables. In addition, there was no significant main effect of experiment and interaction between group and experiment for age, all the OCI-R sub-scales and DASS sub-scales (Table 5.2).

These findings showed that the participant selection process worked well for these experiments and that groups can be compared across experiments. Sampling differences between experiments cannot explain behavioural measures effects.
Table 5.2.

Mean, standard error and statistics for the comparison of checkers and non-checkers sample characteristics across the positive-neutral and negative-neutral experiments. F = female, M = male, Pv = videogame player, NPv = no videogame player, Pm = money player, NPm = no money player, ME = main effect, IA = interaction

<table>
<thead>
<tr>
<th>Variables</th>
<th>Positive-neutral</th>
<th>Negative-neutral</th>
<th>Chi-square or Fisher’s exact test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Checkers</td>
<td>Non-checker</td>
<td>Checkers</td>
</tr>
<tr>
<td>Demographic questionnaire (Categorical variables)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>5 F, 6 M</td>
<td>9 F, 2 M</td>
<td>10 F, 2 M</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Play videogame</td>
<td>5 Pv, 6 NPv</td>
<td>6 Pv, 5 NPv</td>
<td>6 Pv, 6 NPv</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gambling for money</td>
<td>2 Pm, 9 NPm</td>
<td>2 Pm, 9 NPm</td>
<td>1 Pm, 11 NPm</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographic questionnaire (continuous variables)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>20.36 ± 0.60</td>
<td>22.00 ± 1.15</td>
<td>20.25 ± 0.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### OCI-R scores

<table>
<thead>
<tr>
<th>OCI-R Behavior</th>
<th>Mean 1</th>
<th>SD 1</th>
<th>Mean 2</th>
<th>SD 2</th>
<th>F(1, 42)</th>
<th>p</th>
<th>F(1, 42)</th>
<th>p</th>
<th>F(1, 42)</th>
<th>p</th>
<th>η²p &lt;group&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCI-R Checking</td>
<td>7.36±0.59</td>
<td>0.81±0.22</td>
<td>7.25±0.68</td>
<td>0.91±0.25</td>
<td>F(1, 42) = 0,</td>
<td>p = .98</td>
<td>F(1, 42) = 173.30,</td>
<td>p &lt; .001**</td>
<td>F(1, 42) = 0.04,</td>
<td>p = .82</td>
<td>η²p &lt;group&gt; = .80</td>
</tr>
<tr>
<td>OCI-R Washing</td>
<td>2.36±0.62</td>
<td>1.18±0.26</td>
<td>2.66±0.72</td>
<td>1.08±0.68</td>
<td>F(1, 42) = 0.02,</td>
<td>p = .86</td>
<td>F(1, 42) = 5.18,</td>
<td>p = .02*</td>
<td>F(1, 42) = 0.10,</td>
<td>p = .74</td>
<td>η²p &lt;group&gt; = .11</td>
</tr>
<tr>
<td>OCI-R Hoarding</td>
<td>4.81±0.68</td>
<td>2.36±0.57</td>
<td>4.50±0.67</td>
<td>1.50±0.43</td>
<td>F(1, 42) = 0.87,</td>
<td>p = .35</td>
<td>F(1, 42) = 18.65,</td>
<td>p &lt; .001**</td>
<td>F(1, 42) = 0.18,</td>
<td>p = .66</td>
<td>η²p &lt;group&gt; = .30</td>
</tr>
<tr>
<td>OCI-R Ordering</td>
<td>6.27±1</td>
<td>2.81±0.60</td>
<td>5.41±0.60</td>
<td>2.58±0.63</td>
<td>F(1, 42) = 0.56,</td>
<td>p = .45</td>
<td>F(1, 42) = 18.78,</td>
<td>p &lt; .001**</td>
<td>F(1, 42) = 0.18,</td>
<td>p = .67</td>
<td>η²p &lt;group&gt; = .30</td>
</tr>
<tr>
<td>OCI-R Counting</td>
<td>2.63±0.80</td>
<td>0.27±0.19</td>
<td>3.91±0.91</td>
<td>0.33±0.14</td>
<td>F(1, 42) = 1.15,</td>
<td>p = .29</td>
<td>F(1, 42) = 22.63,</td>
<td>p &lt; .001**</td>
<td>F(1, 42) = 0.95,</td>
<td>p = .33</td>
<td>η²p &lt;group&gt; = .35</td>
</tr>
<tr>
<td>OCI-R Obsessing</td>
<td>3.63±0.52</td>
<td>1.00±0.30</td>
<td>3.91±0.89</td>
<td>1.16±0.42</td>
<td>F(1, 42) = 0.14,</td>
<td>p = .70</td>
<td>F(1, 42) = 20.67,</td>
<td>p &lt; .001**</td>
<td>F(1, 42) = 0.009,</td>
<td>p = .92</td>
<td>η²p &lt;group&gt; = .33</td>
</tr>
<tr>
<td>OCI-R TOTAL</td>
<td>27.09±2.30</td>
<td>8.45±1.32</td>
<td>27.66±3.16</td>
<td>7.58±1.25</td>
<td>F(1, 42) = 0.005,</td>
<td>p = .94</td>
<td>F(1, 42) = 78.11,</td>
<td>p &lt; .001**</td>
<td>F(1, 42) = 0.10,</td>
<td>p = .74</td>
<td>η²p &lt;group&gt; = .65</td>
</tr>
</tbody>
</table>

### DASS scores

<table>
<thead>
<tr>
<th>DASS subscale</th>
<th>Mean 1</th>
<th>SD 1</th>
<th>Mean 2</th>
<th>SD 2</th>
<th>F(1, 42)</th>
<th>p</th>
<th>F(1, 42)</th>
<th>p</th>
<th>F(1, 42)</th>
<th>p</th>
<th>η²p &lt;group&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>DASS Depression</td>
<td>5.72±1.33</td>
<td>2.63±0.85</td>
<td>6.08±1.29</td>
<td>2.25±1.05</td>
<td>F(1, 42) = 0,</td>
<td>p = .98</td>
<td>F(1, 42) = 9.38,</td>
<td>p = .004</td>
<td>F(1, 42) = 0.10,</td>
<td>p = .74</td>
<td>η²p &lt;group&gt; = .18</td>
</tr>
<tr>
<td>DASS Anxiety</td>
<td>6.27±1.28</td>
<td>1.09±0.34</td>
<td>6.08±1.29</td>
<td>2.41±0.74</td>
<td>F(1, 42) = 0.31,</td>
<td>p = .57</td>
<td>F(1, 42) = 19.3,</td>
<td>p &lt; .001</td>
<td>F(1, 42) = 0.56,</td>
<td>p = .45</td>
<td>η²p &lt;group&gt; = .31</td>
</tr>
<tr>
<td>DASS Stress</td>
<td>9.18±1.67</td>
<td>3.54±0.74</td>
<td>11.16±1.67</td>
<td>3.50±1.06</td>
<td>F(1, 42) = 0.51,</td>
<td>p = .47</td>
<td>F(1, 42) = 24.04,</td>
<td>p &lt; .001</td>
<td>F(1, 42) = 0.56,</td>
<td>p = .45</td>
<td>η²p &lt;group&gt; = .36</td>
</tr>
</tbody>
</table>
5.2.4. Laboratory-based Study

5.2.4.1. Laboratory-based study procedure

Upon arrival in the laboratory, participants were randomly assigned to one of the two binary decision making task experiments, i.e. the positive-neutral experiment (study 2) and the negative-neutral experiment (study 3). Afterwards, they received the Participant Information Sheet (PIS), either containing information about the positive-neutral experiment or about the negative-neutral experiment (Appendix 2.4 for an example of PIS), and signed a written consent form (Appendix 2.5). The duration of the entire laboratory-based study was approximately 2 hours. It contained the standard Picture Validation task, a binary decision making task, and the standard laboratory-based questionnaires which were the Domain Specific Risk-Taking Scale (DOSPERT); Intolerance of uncertainty questionnaire (IUS); Impulsive Behaviour Scale (UPPS) and the South-Oaks Gambling Screen. Participants also completed a strategy questionnaires (see Chapter 2 for methods details).

The binary decision making tasks were as described in Chapter 2. They only differed in their feedback type. The tasks were designed to investigate the effect of feedback direction (positive vs. negative) on RBL. The specific feedback types for the two experiments were positive monetary feedback and neutral pictures for the positive-neutral experiment (study 2) and negative monetary feedback and neutral pictures for the negative-neutral experiment (study 3). For these experiments, 40 neutral pictures (previously selected in Chapter 3) were divided in two sets of 20 neutral pictures. These neutral picture sets were matched for valence (set 1: 4.56±0.10; set 2: 4.31±0.11; t(38) = 1.56, p=.12). They were counterbalanced for target-positions and target probabilities in each experiment.

To investigate if a particular neutral picture set was differently affecting decision making in groups, one set was employed in half of the experiments. Half of participants performed an experiment containing set of pictures 1 and the other half performed the experiment with the set of pictures 2.

5.2.4.2. Data analysis for the laboratory-based study

For the picture validation task, data analysis was conducted separately for the measures of valence, arousal and comfort ratings. For each measure, a three-way 4 x 2 x 2 mixed ANOVA with the within-participant factor Picture Type (positive, neutral, negative
and checking-related pictures) and the between-participant factors Group (checkers vs. non-checkers) and Experiment (positive-neutral vs. negative-neutral) was conducted. Post-hoc t-tests were performed and Bonferroni corrections applied when necessary.

For the binary decision making task, accuracy, sub-block accuracy, exploration/ exploitation (win shift/lose shift) and cross-correlation measures were analysed. These analysis are described in detail in Chapter 2. Table 5.3 gives an overview of the ANOVA types conducted for each of these measures. Generally, blocks were analysed separately because of their maximum accuracy differences and their structural differences. For all measures, all three blocks were analysed, except for the exploration/exploitation (win shift/lose shift measure) where only block 1 was analysed.

Table 5.3.
Statistical ANOVA analyses conducted on the measures of the binary decision making task in this study.

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Within-participant factor</th>
<th>Between-participant factor</th>
<th>Type of test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Group</td>
<td>Experiment</td>
</tr>
<tr>
<td>Accuracy</td>
<td>None</td>
<td>Checkers vs. non-checkers</td>
<td>Two-way 2 x 2 ANOVA</td>
</tr>
<tr>
<td>Sub-block accuracy</td>
<td>5 Sub-blocks</td>
<td>Positive-neutral vs. negative-neutral</td>
<td>Three-way 5 x 2 x 2 mixed-ANOVA</td>
</tr>
<tr>
<td>Win-shift</td>
<td>None</td>
<td>Checkers vs. non-checkers</td>
<td>Two-way 2 x 2 ANOVA</td>
</tr>
<tr>
<td>Lose-shift</td>
<td>none</td>
<td>Checkers vs. non-checkers</td>
<td>Two-way 2 x 2 ANOVA</td>
</tr>
<tr>
<td>Cross-correlation for each block</td>
<td>5 Lags</td>
<td>Positive-neutral vs. negative-neutral</td>
<td>Three-way 5 x 2 x 2 mixed-ANOVA</td>
</tr>
</tbody>
</table>

The lab-based questionnaires were compared between groups and experiments using 2 x 2 ANOVAs with the factors group (checker vs. controls) and Experiment (positive-neutral experiment vs. negative neutral experiment). The exception was the strategy questionnaires
that was analysed by using a Pearson Chi-square or Fisher’s exact test to compare groups for each experiment and experiments for each group. No interaction effects were investigated for the strategy questionnaire. Findings for these lab-based questionnaires will be discussed together with Chapter 6 and 7 findings in Chapter 8 (General discussion), except for the strategy questionnaires that will be also discussed here in Chapter 5.4.

Finally, Kendall-tau correlations were computed for each experiment separately, considering only checkers to investigate if (1) symptom strength recorded with OCI-R total and OCI-R checking subscale scores, (2) comorbidities recorded with the DASS, and (3) measures from the laboratory-based questionnaires are related to behavioural measures recorded in the binary decision making task. These measures are: Accuracies for block 1, block 2, and block 3; win-shift and lose-shift; cross-correlations for lags 1-5 for block 1, 2 and 3.

5.3. Results

5.3.1. Picture Validation task

5.3.1.1. Valence scale

The mixed 4 x 2 x 2 ANOVA with the factors Picture Type, Group and Experiment revealed a significant main effect of Picture Type, \(F(2,74,106.79) = 414.92, p < .001, η^2_p = .90\). The related post-hoc paired t-tests were all statistically significant, all \(t(45) > 3.50, all p’s < .01\). As expected from previous results in Chapter 3, positive pictures were rated as more positive than neutral, checker-related and negative pictures; neutral pictures were rated as neutral and more positive than checker-related and negative pictures and checker-related pictures were rated as less negative than negative pictures.

In addition, a marginal interaction between the factors Picture Type and Group was present, \(F(2,74,106.79) = 2.71, p = .058, η^2_p = .06\). Four independent post-hoc t-tests were computed to compare groups for each picture type across both experiments. The analysis showed no significant difference between groups for any of the picture types, including the checking-related pictures ((checkers: \(4.98±0.25\); non-checkers: \(4.51±0.12\); \(t(31.97) = 1.63, p=.11\)). No other effect was significant.
5.3.1.2. Arousal scale

ANOVA results showed only a significant main effect of Picture Type, $F(1.71,72.08)=18.81, p < .001$. To investigate the significant main effect of picture type, 6 pairwise comparisons were conducted. Not all pairwise comparisons were statistically significant and they are shown in table 5.4. Basically, they showed that positive, negative and checker-related pictures generated more arousal than neutral pictures. Positive and negative pictures and positive and checker-related pictures presented similar amounts of arousal. Checker-related pictures, however, presented less arousal than negative pictures, considering both experiments and groups.

Table 5.4.  
Pairwise comparisons for picture type in the picture validation task in the arousal scale.  
(Lower values of arousal indicate more arousal)

<table>
<thead>
<tr>
<th>Picture-type (Arousal scale)</th>
<th>Mean and standard error</th>
<th>T-test</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive vs. Neutral</td>
<td>4.02±0.22 vs. 5.42±0.12</td>
<td>t(45) = -6.67, $p &lt; .001^{**}$, $r = .70$</td>
<td></td>
</tr>
<tr>
<td>Positive vs. Negative</td>
<td>4.02±0.22 vs. 4.06±0.19</td>
<td>t(45) = -0.12, $p = .89$</td>
<td></td>
</tr>
<tr>
<td>Positive vs. Checker-related</td>
<td>4.02±0.22 vs. 4.59±0.15</td>
<td>t(45) = -2.22, $p = .18$</td>
<td></td>
</tr>
<tr>
<td>Neutral vs. Negative</td>
<td>5.42±0.12 vs. 4.06±0.19</td>
<td>t(45) = 7.47, $p &lt; .001^{**}$, $r = .74$</td>
<td></td>
</tr>
<tr>
<td>Neutral vs. Checker-related</td>
<td>5.42±0.12 vs. 4.59±0.15</td>
<td>t(45) = 6.52, $p &lt; .001^{**}$, $r = .69$</td>
<td></td>
</tr>
<tr>
<td>Negative vs. Checker-related</td>
<td>4.06±0.19 vs. 4.59±0.15</td>
<td>t(45) = -4.06, $p &lt; .001^{**}$, $r = .51$</td>
<td></td>
</tr>
</tbody>
</table>

5.3.1.3. Comfort scale

Results showed there was a significant main effect of Picture Type, $F(2.98,125.39) = 285.22, p < .001, \eta^2 = .87$. All post-hoc paired t-tests were statistically significant, all $t(45) > 3.50$, all $p’s < .01$. Basically, positive pictures were rated as more comfortable than neutral, checker-related and negative pictures; neutral pictures were rated as more comfortable than
negative and checker-related pictures and checker-related pictures as more comfortable than negative pictures.

In addition, there was a significant interaction between the factors Picture Type and Group, $F(2.98, 125.39) = 3.31, p = .02, \eta^2 = .07$. The post-hoc t-test analysis did not reveal any significant group differences for the positive, neutral and negative pictures. However, a significant group difference was found for the checking-related pictures, (Mean (Checker): $6.23 \pm 0.22$; Mean (Non-checker): $5.30 \pm 0.17$; $t(44) = 3.24, p_{b} = .008, r = .55$), showing that checkers were more uncomfortable with the checking-related pictures than non-checkers (Figure 5.2). No other effect was significant in this analysis.

**Figure 5.2** Mean comfort ratings for positive, neutral, negative, and checking-related pictures displayed separately for checkers and non-checkers. Lower scores represent more positive comfort and higher scores mean more negative comfort. Neutral ratings have a score close to 5. Error bars represent standard error measures.

In summary, the findings from the picture validation task study reported in Chapter 3 were only partly confirmed in this study. In Chapter 3, group differences were reported for the valence ratings of the checking-related pictures and the comfort ratings of the negative and checking-related pictures. However, in Chapter 5, group differences were only found for the comfort ratings of the checking-related pictures. These rating differences between studies will be discussed in Chapter 8 (General Discussion).
5.3.2. Binary decision making task

5.3.2.1. Accuracy

Mean accuracy values for each block, group and experiment are displayed in Figure 5.3. For each block, a two-way between-participant 2 x 2 ANOVA with the factors Group (checker vs. non-checker) and Experiment (study 2 as positive-neutral vs. study 3 as negative neutral experiment) was conducted to investigate group and experiment differences for each block.

The analysis of block 1 revealed a significant main effect of experiment for accuracy in block 1, $F(1,42) = 4.40, p = .04, \eta^2_p = .09$, showing that the mean accuracy was higher for the negative-neutral experiment (0.68 ± 0.01) compared to the positive-neutral experiment (0.56± 0.01).

For block 2, there was a significant interaction between the factors Group and Experiment, $F(1,42) = 9.87, p = .003, \eta^2_p = .19$. Post-hoc independent t-tests showed that checkers (0.68 ± 0.03) had significantly higher accuracy levels than non-checkers (0.55± 0.03; $t(22) = 2.61, p_b = .03, r = .48$) in the negative-neutral experiment, whereas no significant group differences were found for the positive-neutral experiment (Mean (Checker): 0.57 ± 0.03; Mean (Non-checker): 0.65± 0.02; $t(20) = 1.82, p_b = .16$) (Figure 5.3). Secondly, checkers had significantly higher accuracy scores in the negative-neutral experiment compared to the positive-neutral experiment (Mean (Negative-neutral): 0.68 ± 0.03; Mean (Positive-neutral): 0.57± 0.02; $t(21) = 2.43, p_b = .05$), whereas non-checkers had no significant accuracy differences between experiments (Mean (Negative-neutral): 0.56 ± 0.03; Mean (Positive-neutral): 0.65± 0.02; $t(21) = 2.01, p_b = .14$). (Figure 5.3).

For block 3, there were no significant main effects of group and experiment and no interaction between both factors.
Figure 5.3 Mean accuracy values for checkers and non-checkers in the negative-neutral and positive-neutral experiment for blocks 1, 2 and 3.

5.3.2.2. Sub-block Accuracy (Learning curves)

A step further was taken considering the initial accuracy analysis by looking at sub-blocks to explore accuracy changes over the duration of a block. i.e. learning curves. These learning curves are displayed in Figure 5.4 and it seems that accuracy improves over time for all blocks. We also investigated group and experimental differences in these learning curves.
For this, a three-way 5 x 2 x 2 mixed ANOVA was conducted for each block. There were significant main effects of Sub-block for block 1 (\(F(4,168) = 11.09, p < .001, \eta^2 = .20\)), block 2 (\(F(4,168) = 7.52, p < .001, \eta^2 = .15\)), and block 3 (\(F(4,168) = 6.65, p < .001, \eta^2 = .13\)). Post-hoc paired t-tests are displayed in Table 5.5.
Table 5.5

Pairwise comparisons for the main effect of sub-block accuracy for blocks 1, 2 and 3

<table>
<thead>
<tr>
<th>Sub-block accuracy</th>
<th>Mean and standard error</th>
<th>T-test</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Block 1</td>
<td>Block 2</td>
<td>Block 3</td>
</tr>
<tr>
<td>Sub-block 1 vs. Sub-block 2</td>
<td>0.50±0.01 vs. 0.54±0.01</td>
<td>t(45) = -1.83, pb</td>
<td>r = .61</td>
</tr>
<tr>
<td>Sub-block 1 vs. Sub-block 3</td>
<td>0.50±0.01 vs. 0.56±0.01</td>
<td>t(45) = -3.05, pb</td>
<td>r = .61</td>
</tr>
<tr>
<td>Sub-block 1 vs. Sub-block 4</td>
<td>0.50±0.01 vs. 0.58±0.01</td>
<td>t(45) = -3.78, pb</td>
<td>r = .61</td>
</tr>
<tr>
<td>Sub-block 1 vs. Sub-block 5</td>
<td>0.50±0.01 vs. 0.62±0.01</td>
<td>t(45) = -6.29, pb</td>
<td>r = .61</td>
</tr>
<tr>
<td>Sub-block 2 vs. Sub-block 3</td>
<td>0.54±0.01 vs. 0.56±0.01</td>
<td>t(45) = -1.92, pb</td>
<td>r = .61</td>
</tr>
<tr>
<td>Sub-block 2 vs. Sub-block 4</td>
<td>0.54±0.01 vs. 0.58±0.01</td>
<td>t(45) = -4.31, pb</td>
<td>r = .61</td>
</tr>
<tr>
<td>Sub-block 2 vs. Sub-block 5</td>
<td>0.54±0.01 vs. 0.62±0.01</td>
<td>t(45) = -4.31, pb</td>
<td>r = .61</td>
</tr>
<tr>
<td>Sub-block 3 vs. Sub-block 4</td>
<td>0.56±0.01 vs. 0.58±0.01</td>
<td>t(45) = -0.94, p</td>
<td>r = .61</td>
</tr>
<tr>
<td>Sub-block 3 vs. Sub-block 5</td>
<td>0.56±0.01 vs. 0.62±0.01</td>
<td>t(45) = -3.39, pb</td>
<td>r = .61</td>
</tr>
<tr>
<td>Sub-block 4 vs. Sub-block 5</td>
<td>0.56±0.01 vs. 0.62±0.01</td>
<td>t(45) = -2.31, pb</td>
<td>r = .61</td>
</tr>
</tbody>
</table>
These post-hoc tests revealed that participants were able to increase performance in all the three blocks; as accuracy levels progressively increased across sub-blocks. However, learning in block 1 was different from blocks 2 and 3. In block 1, participants needed at least two sub-blocks to increase accuracy (from sub-blocks 1 to 3 and again from 3 to 5), whilst they only needed the first sub-block (from sub-block 1 to 2) to learn the task in blocks 2 and 3 (see Figure 5.4), which could be related to the difficulty levels of each block but also to the fixed block order.

There was no others significant main effects, no significant interaction between experiment and sub-block accuracy for blocks 1, 2 and 3, no significant interaction between group and sub-block accuracy for blocks 1, 2 and 3 and no significant interaction between experiment, group and sub-block for all the blocks, except for block 2, where there was a marginally significant interaction, $F(4.,168) = 2.05, p = .08, \eta^2_p = .04$ of experiment vs. group vs. sub-block.

Because the previous accuracy analysis (see Chapter 5.3.2.1.) revealed an interaction between the factors Group and Experiment, this marginal effect was further investigated with two-way 2 x 2 ANOVAs for each sub-block separately. These post-hoc ANOVAs revealed a significant main effect of experiment for sub-block 2, $F(1.,42) = 4.30, p = .04, \eta^2_p = .09$, meaning that participants performed better in sub-block 2 of the negative-neutral experiment compared to sub-block 2 of the positive-neutral experiment (Mean (negative-neutral): 0.65 ± 0.02; Mean (positive-neutral): 0.58 ± 0.02). More importantly, significant interactions between the factors Group and Experiment were found for the sub-blocks 2-5 (table 5.6)

<table>
<thead>
<tr>
<th>Sub-block</th>
<th>Statistics</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>$F(1.,42) = 7.34, p = .01^*$</td>
<td>$\eta^2_p = .14$</td>
</tr>
<tr>
<td>3</td>
<td>$F(1.,42) = 8.65, p = .005^{**}$</td>
<td>$\eta^2_p = .17$</td>
</tr>
<tr>
<td>4</td>
<td>$F(1.,42) = 7.72, p = .008^{**}$</td>
<td>$\eta^2_p = .15$</td>
</tr>
<tr>
<td>5</td>
<td>$F(1.,42) = 7.30, p = .01^*$</td>
<td>$\eta^2_p = .14$</td>
</tr>
</tbody>
</table>

For each of these interactions, independent t-tests were conducted for each experiment between groups and for each group between experiments. For the negative-neutral experiment, significantly higher accuracy scores were recorded for checkers compared to
non-checkers in sub-blocks 2, 3 and 5 (Table 5.7; Figure 5.5), whereas no group accuracy differences were found for the sub-blocks of the positive-neutral experiment (Table 5.7). In the analysis of each group between experiments, there was no significant effect of experiment for checkers for sub-blocks 3 to 5 but there was a significant effect of experiment in sub-block 2, where checkers in the negative-neutral experiment presented a higher accuracy than checkers in the positive-neutral experiment, ((Mean (negative-neutral): 0.71 ± 0.02; Mean (positive-neutral): 0.54± 0.03; \( t(21) = 3.93 \), \( p_b = .02 \), \( r=0.65 \)). For non-checkers, there was no significant effect of experiment for all the concerned sub-blocks (Figure 5.5).

Table 5.7

Independent post-hoc t-tests for each experiment between groups with mean accuracy levels and standard error values

<table>
<thead>
<tr>
<th>Sub-blocks</th>
<th>Negative-neutral experiment</th>
<th>Positive-neutral experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Checkers</td>
<td>Non-checkers</td>
</tr>
<tr>
<td>Block 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.71±0.02</td>
<td>0.60±0.03</td>
</tr>
<tr>
<td>3</td>
<td>0.73±0.05</td>
<td>0.55±0.04</td>
</tr>
<tr>
<td>4</td>
<td>0.71±0.03</td>
<td>0.56±0.05</td>
</tr>
<tr>
<td>5</td>
<td>0.74±0.04</td>
<td>0.57±0.04</td>
</tr>
</tbody>
</table>
Figure 5.5 Sub-block accuracy mean values for checkers and non-checkers, divided between experiments for block 2.

In summary, the accuracy analysis for each block separately revealed, firstly, that learning occurred in all three blocks. Note that it was slower in block 1, however. Secondly, there were no group or experimental differences for the random sequence block (block 1). Thirdly, checkers learned the pattern sequences in block 2 better when negative monetary feedback was provided compared to the positive monetary feedback. In accordance, checkers learned better than non-checkers in the negative-neutral experiment. These differences were only present from sub-block 2 onwards (after 32 trials). Finally, no group or experimental differences were found for block 3 where target option maximisation and pattern search strategies were combined. These findings will be discussed later on.

5.3.2.3. Exploration / exploitation: Win-stay / lose-shift measures in Block 1

Results from the two two-way $2 \times 2$ ANOVA showed there was no significant main effects of Experiment or Group, and no interaction between the factors Experiment and Group for win-shift and for lose-shift measures. These measures were not affected by the feedback direction manipulation and were not different for subclinical checkers and non-
checkers. Figure 5.6 shows win-shift / lose-shift strategies per participants for checkers and non-checkers and for each experiment. This figure indicates that checkers and non-checkers are not doing excessive exploration (more win-shift and more lose-shift) because no participant presented a higher probability of shifting after losing and winning in the positive-neutral and in the negative-neutral experiments (see right superior quadrant for both experiments in figure 5.6). Besides, it is possible to see that both groups prefer to shift after losing, as only one participant is seen at the right inferior quadrant of figure 5.6 (more win-shift/lose-stay), where the probability of shifting after losing is the lowest in the positive-neutral and in the negative-neutral experiments. Additionally, both groups seem to not be doing excessive exploitation, as their probability of shifting after losing and winning seem to be about 0.6 and not at the lowest values of shifting (see left inferior quadrant). However, it is interesting to note that checkers seem to be doing more shifts after losing than non-checkers in the positive-neutral experiment, even though the difference between groups is not significant for the probability of shifting after losing.

**Figure 5.6** Win-shift / Lose-shift strategies per participant divided by group and experiment. The left and inferior quadrant of the figure indicates *more exploitation after winning and after losing* (win-stay / lose-stay). The right and superior quadrant indicate *more exploration after winning and losing* (win-shift / lose-shift). The left and superior quadrant indicates *more*
exploration after losing only (win-stay / lose-shift). The right and inferior quadrant indicates more exploration after winning only (win-shift / lose-stay)

5.3.2.4. Cross-correlations

5.3.2.4.1. Cross-correlations for Block 1

Figure 5.7. shows the experimental auto-correlation- and cross-correlations for block 1. Here, the difference between auto-and cross-correlation scores seems to be most evident for lag 1, but maybe also present for lag 3.

![Cross-correlation graph](image)

*Figure 5.7* Comparison between the auto-correlation values (experimental sequence) and the cross-correlation values (participant) for lags 1 to 5 in block 1. Values were averaged across groups and experiments.

Actually, the three-way 5 x 2 x 2 mixed ANOVA on the auto- and cross-correlation difference scores revealed a significant main effect of lag, $F(3.85,.161.72) = 13.68, p < .001$, $\eta^2 = .24$. This was caused by a significantly higher difference between the participant’s auto- and cross-correlation values for lag 1 than for lags 2, 3, 4, and 5, all $t(45) > 3.50$, all $p’s < .01$. This might be because participants assumed a greater dependency between the current and the previous event (lag 1) than it was supposed for a random sequence as in block 1. More specifically, participants used the previous trial to make decisions, thereby, increasing the cross-correlation value for lag 1. In fact, this also augmented the difference between the participants and experimental sequence correlation value. In addition, the difference between auto- and cross-correlations was higher at lag 3 compared to lag 2, (Mean (lag 2): -0.02 ±
0.02; Mean (lag 3): 0.07± 0.02; $t(45) = 4.35, p < .001, r=0.54$), an effect that is more difficult to explain. It might be more due to uncontrolled changes in the experimental sequences, because the experimental sequence score is slightly enhanced at lag 3 and because this difference was not found for the others pairwise comparisons of lag 2 (2-4, 2-5) or lag 3 (3-4, 3-5). No other main effects and interactions were significant and, hence, no group differences were found for block 1.

### 5.3.2.4.2. Cross-correlations for Block 2

Figure 5.8 shows that both groups generally followed the pattern in the experimental sequence. In this way, checkers and non-checkers were able to learn sequential effects in this block. As expected the absolute cross-correlation values were not as large as the auto-correlation values. Figure 5.8 also reveals group and experimental differences in the participant responses to the pattern sequence where checkers performing in the negative-neutral experiment seem to be closest to the auto-correlation values than non-checkers and closest in the negative-neutral experiment in comparison with the positive-neutral experiment.

![Cross-correlation probability from lags 1 to 5 between checkers and non-checkers for the negative-neutral and positive-neutral experiment in block 2. The averaged correlation values of the experimental sequence are shown in black.](image)

*Figure 5.8* Cross-correlation probability from lags 1 to 5 between checkers and non-checkers for the negative-neutral and positive-neutral experiment in block 2. The averaged correlation values of the experimental sequence are shown in black.
Actually, a three-way ANOVA 5 x 2 x 2 for block 2 showed there was a significant interaction of lags vs. experiment vs. group, $F(1.72, .72.50) = 8.03, p < .001, \eta^2_p = .16$ and there was a main effect of lag, $F(1.72, .72.50) = 72.53, p < .001, \eta^2_p = .63$, that was expected as lag values were different in the sequence, due to the presence of sequential effects, linking specific previous trials to the actual trial. Additionally, no significant effects of lags vs. experiment and lags vs. group were found.

In fact, the significant interaction was furtherly explored with five two-way 2 x 2 ANOVA from lags 1 to 5 with the factors Experiment and Group. Results showed a significant interaction between experiment and group for lag 2, $F(1, 42) = 10.77, p < .002, \eta^2_p = .20$, lag 4, $F(1, 42) = 9.78, p < .003, \eta^2_p = .18$ and a marginal effect for lag 5, $F(1, 42) = 3.86, p = .056, \eta^2_p = .08$. For lags 2 and 4, these interactions were caused by lower differences between auto-and cross-correlations for checkers in the negative neutral experiments compared to the positive-neutral experiment, and for checkers compared to non-checkers in the negative-neutral experiment (table 5.11), indicating checkers were closer to the original sequence and, therefore, better pattern learners. This data corroborates with the accuracy findings for block 2. For lag 5, there was no significant effect of group but non-checkers presented lower differences between auto-and cross-correlations in the positive-neutral experiment than in the negative-neutral experiment, indicating non-checkers were closer to the experimental sequence in lag 5 than checkers, consequently suggesting that non-checkers were closer to the original sequence in the positive-neutral than in the negative-neutral experiment in block 2., therefore better pattern learners in this feedback direction. Detailed post-hoc t-tests are presented in Table 5.8.
Table 5.8.  
Independent t-tests for lags 2, 4 and 5 for block 2 between groups for both experiments

<table>
<thead>
<tr>
<th>Lags</th>
<th>Negative-neutral</th>
<th>Positive-neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Checkers</td>
<td>Non-checkers</td>
</tr>
<tr>
<td>Lag 2</td>
<td>-0.20±0.08</td>
<td>-0.57±0.09</td>
</tr>
<tr>
<td></td>
<td>( t(22) = 2.89, p_b = .02^* ), ( r = 0.52 )</td>
<td>( t(22) = 1.73, p = .09 )</td>
</tr>
<tr>
<td>Lag 4</td>
<td>0.13±0.06</td>
<td>0.39±0.07</td>
</tr>
<tr>
<td></td>
<td>( t(22) = 2.60, p_b = .03^* ), ( r = 0.48 )</td>
<td>( t(20) = 1.81, p = .08 )</td>
</tr>
<tr>
<td>Lag 5</td>
<td>-0.14±0.03</td>
<td>0.18±0.03</td>
</tr>
<tr>
<td></td>
<td>( t(22) = 0.84, p = .40 )</td>
<td>( t(20) = 2.00, p_b = .12 )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lags</th>
<th>Negative-neutral</th>
<th>Positive-neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Checkers</td>
<td>Non-checkers</td>
</tr>
<tr>
<td>Lag 2</td>
<td>-0.20±0.08</td>
<td>-0.57±0.09</td>
</tr>
<tr>
<td></td>
<td>( t(21) = 2.96, p_b = .01^* ), ( r = 0.54 )</td>
<td>( t(21) = 1.71, p = .10 )</td>
</tr>
<tr>
<td>Lag 4</td>
<td>0.13±0.06</td>
<td>0.39±0.07</td>
</tr>
<tr>
<td></td>
<td>( t(21) = 3.04, p_b = .01^* ), ( r = 0.55 )</td>
<td>( t(21) = 1.37, p = .18 )</td>
</tr>
<tr>
<td>Lag 5</td>
<td>-0.14±0.03</td>
<td>0.15±0.04</td>
</tr>
<tr>
<td></td>
<td>( t(21) = 0.24, p = .80 )</td>
<td>( t(21) = 2.77, p_b = .02^* )</td>
</tr>
</tbody>
</table>

**5.3.2.4.3. Cross-correlations for Block 3**

The three-way ANOVA 5 x 2 x 2 for block 3 showed a main effect of Lag, \( F(1.72,72.50) = 57.52, p < .001, \eta^2 = .57 \), that was expected as lag values were different in the sequence, due to the presence of sequential effects. This effect was not further investigated. In addition, a marginal interaction between the factors Lag, Experiment, and Group was found, \( F(1.73,72.77) = 2.89, p = .06, \eta^2 = .06 \). It was expected that this effect might be caused by group and experimental differences in lag 3. Hence, five two-way 2 x 2 ANOVA with the factors Experiment and Group were conducted. However, the analysis did not reveal any significant effects (Figure 5.9).
Figure 5.9 Cross-correlation probability from lags 1 to 5 between checkers and non-checkers for the negative-neutral and positive-neutral experiment in block 3. The averaged correlation values of the experimental sequence are shown in black.

In summary, the findings from the auto- and cross-correlation analysis are in accordance with the results from the accuracy analysis. No group or experimental differences were found for block 1 (random block with target option probability manipulation) and block 3 (target option and conditional probability manipulation). However, group differences were revealed for block 2. More specifically, checkers learned the pattern in block 2 better in the negative-neutral experiment compared to the positive neutral experiment. They were also better compared to non-checkers in the negative-neutral experiment. In addition, non-checkers learned the pattern in block 2 better in the positive-neutral experiment compared to the negative-neutral experiment. These findings will be discussed below.

5.3.3. Laboratory-based questionnaires

5.3.3.1. Intolerance of Uncertainty Scale (IUS)

The two-way 2 x 2 ANOVA analysis with the factors Experiment and Group showed a significant main effect of group for factor 2 (uncertainty is unfair and spoils everything), $F(1,42) = 5.06, p = .03, \eta^2 = .10$, meaning that checkers (Mean: 36.39 ± 2.17) had higher
levels of intolerance of uncertainty compared to non-checkers (Mean: 29.78 ± 1.83). In addition, there was a marginally significant main effect of group for the total IUS score, (Mean (checkers): 74.13 ± 4.03; Mean (non-checkers): 62.60± 4.03; \(F(1,42) = 3.87, p = .056, \eta^2 = .08\), which is likely to be caused by the enhanced factor 2 ratings for checkers.

5.3.3.2. Domain Specific Risk-Taking Scale and Risk-Perception scale (DOSPERT)

The statistical analysis showed a marginally significant main effect of experiment for financial risk-taking, (Mean (negative-neutral): 13.66 ± 1.31; Mean (positive-neutral): 17.45± 1.37; \(F(1,42) = 3.97, p = .053, \eta^2 = .08\), where participants risked less in the negative-neutral than in the positive-neutral experiment.

5.3.3.3. UPPS Impulsivity Scale

There were no significant main effects of experiment, group and interaction for all the UPPS sub-sections and total score. Thus, it seems checkers and non-checkers had similar impulsivity ratings in both experiments.

5.3.3.4. South Oaks Gambling Screen

There was a significant main effect of group for gambling, (Mean (checkers): 1.00 ± 0.29; Mean (non-checkers): 0.21± 0.10; \(F(1,42) = 6.94, p = .012, \eta^2 = .14\), where checkers presented more gambling behaviour than non-checkers.

5.3.3.5. Strategy questionnaires

Categorical data from the strategy questionnaire showed that checkers and non-checkers seem to employ strategies and to visualize patterns in similar ways in both experiments and their respective blocks. The exception was block 2. Here, more checkers saw a pattern in the negative-neutral experiment (11 yes, 1 no) than in the positive-neutral experiment (4 yes, 7 no); Fisher’s exact test, \(p = .009\). Additionally, there was a marginal effect of group for the visualization of a pattern in the negative-neutral experiment, where checkers saw more patterns (11 yes, 1 no) than non-checkers (6 yes, 6 no), \(p = .07\).

In summary, the analysis of the lab-based questionnaires revealed that checkers showed greater levels of intolerance of uncertainty but were also bigger gamblers than non-checkers. In addition, in accordance with the findings from the behavioural measures,
checkers saw more patterns in block 2 of the negative-neutral experiment than in the positive-neutral experiment.

5.3.4. Correlation analysis between behavioural measures from the binary decision making task and laboratory-based questionnaire scores

5.3.4.1. Correlation analysis for the negative-neutral experiment (study 3)

Firstly, Kendall-tau correlations were conducted to investigate whether symptom severity in checkers affected performance in the binary decision making task. The OCI-R checking scores correlated with the pattern learning in block 2, i.e. the correlation with Dif (lag 2) was marginally significant, \( r_\tau = -0.04, p = 0.056 \), and the correlation with Dif (lag 4) was significant, \( r_\tau = 0.46, p = 0.04 \). As the Dif for cross-correlation is calculated considering first the auto-correlation value (AC) of the experimental sequence for one given lag minus the cross-correlation value (CC) of the participant’s sequence for one given lag, this means that lags with positive auto-correlation values, as lag 4 will present lower Difs when participant’s cross-correlation values increase (i.e., Dif (lag 4) = AC – CC). In this way, the lower the Dif value for lag 4, the more cross-correlation the participant would be doing for this lag. For lag 2, however, the AC value is always negative, so when participants present more cross-correlations with this lag, cross-correlation values for this participant will also be negative, so Dif values will be more positive (i.e., Dif (lag 2) = AC – (-CC)). In this way, the more cross-correlation the participant does with lag 2, the higher the Dif value will be. At the end, these significant correlations mean that checkers with lower checking scores learn the pattern better.

Secondly, no significant correlation between the comorbidity measures depression and anxiety and the binary decision making task measures were found, showing no significant influence of comorbidities on probabilistic learning in negative monetary feedback environments.

Finally, Kendall Tau correlations were performed to investigate if higher ratings in laboratory-based measures, such as intolerance of uncertainty, financial risk taking, and gambling would be related to the task performance in checkers. These laboratory-based measures were selected because their scores differed between the checkers and non-checkers. Intolerance of uncertainty ratings negatively correlated with the accuracy levels in block 2, meaning that higher levels of intolerance of uncertainty are linked to lower levels of block 2
accuracy and reduced pattern learning (Factor 1: $r_\tau = -0.61, p = .006$; Factor 2: $r_\tau = -.47, p = .03$; Total score: $r_\tau = -.54, p = .016$). Accordingly, higher intolerance of uncertainty ratings negatively correlated with lag 2 differences between auto-and cross-correlations in block 2 (Factor 1: $r_\tau = -.54, p = .01$; Factor 2: $r_\tau = -.46, p = .03$; Total score: $r_\tau = -.46, p = .03$). In this way, checker participants with higher intolerance of uncertainty scores had lower accuracy and lower cross-correlation values in block 2.

For financial risk-taking, results showed that those checkers with higher risk-taking finance scores were more likely to have a higher probability of shifting after winning, $r_\tau = .50, p = .02$ and to have a higher Dif (lag 1) for block 1, $r_\tau = .46, p = .03$, caused by the less random behaviour at lag 1 due to the enhanced shifting after winning. No significant correlations were found for the gambling questionnaire.

To summarise the correlation findings for the negative-neutral experiment, high levels of checking and intolerance of uncertainty are related to reduced performance in block 2 of the binary decision making task which is the pattern search block. In addition, financial risk-taking is related to more shifts after winning which is also reflected in more correlations with the cross-correlation scores at lag 1 in block 1.

5.3.4.2. Correlation analysis for the positive-neutral experiment (study 2)

There were no group or experimental differences in the positive neutral experiment. Nevertheless, Kendall Tau correlations were performed to explore the relation between symptom severity and other laboratory-based questionnaire measures and the behavioural measures from the binary decision making task.

Firstly, enhanced symptom severity (OCI-R total) was related to better performance in block 1, $r_\tau = .50, p = .03$, and block 3, $r_\tau = .54, p = .02$. Secondly, comorbidity levels correlated with the binary decision making task performance. More specifically, higher depression levels in checkers were related to enhanced accuracy levels in the binary decision making task in block 1, $r_\tau = .61, p = .01$. Moreover, higher anxiety levels in checkers were related to higher lose-shift scores, $r_\tau = .51, p = .03$. Finally, laboratory-based questionnaire ratings did not correlate with the task performance measures in the binary decision making task.

In summary, it seems that checkers with higher scores of checking and higher scores of intolerance of uncertainty are learning less in block 2 in the negative-neutral experiment. In contrast, they present a better performance in block 1 of the positive-neutral experiment.
Comorbidities only had an effect on the binary decision making task performance of the positive-neutral experiment. Here, depression was related to enhanced performance in block 1 and anxiety was related to enhanced lose shift behaviour. These findings will be discussed below.

5.4. Discussion

The presented study investigated subclinical checkers and matched non-checkers and compared their performance in two separate experiments that manipulated feedback direction in binary decision making tasks disguised as a binary decision making task. One experiment used positive monetary vs. absent feedback (study 2) and the second experiment used negative monetary vs absent feedback (study 3). In both experiments, three blocks were presented, i.e. a random sequence block with a target-option manipulation, a pattern sequence block with equal winning probabilities for both target options, and a pattern sequence block where one target option had a higher chance of winning.

The first aim of this study was to examine feedback direction effects, i.e. positive vs. negative monetary feedback, in people with subclinical checking symptoms and compare them to non-checkers in a random sequence block and to compare our findings with previously reported studies (See Chapter 5.1). Crucial differences between the presented and the previous studies were the between participant manipulation of positive vs. absent monetary feedback or negative vs. absent monetary feedback, and the usage of longer blocks to evaluate effects during learning and after the learning curve has stabilized.

The findings from the accuracy analysis of all three blocks showed that participants were able to learn the task over time. For block 1, accuracy levels increased from sub-block 1 to sub-block 3 and stabilised thereafter. In other words, participants needed two sub-blocks (64 trials) to increase their performance. Learning curves were steeper for blocks 2 and 3. Here, they only needed one sub-block for learning. This could be because of the nature of the block, i.e. block 1 was more difficult than blocks 2 and 3 (lower maximum accuracy), and it did not present a pattern. However, humans seem to be wired to search patterns in nature (Jones & Pashler, 2007), even when there are no patterns at all (Vulkan et al., 2000; Unturbe & Corominas, 2007) which might explain why it might take longer for participants to stabilise their learning in a random sequence block.

More importantly, no group or experimental accuracy differences were found for the random sequence block 1 with differential target option probabilities. The findings from the
auto- and cross-correlation analysis were in accordance with the results from the accuracy analysis. There were also no group and experimental differences. Only a main effect of lags was found for block 1, indicating that participants were being influenced by the immediately preceding trial to make decisions.

Generally, this data indicate that subclinical checkers are as equally able to learn the random sequence task as non-checkers in both the positive-neutral and the negative-neutral experiment. This seems to partly contradict findings from Starcke et al. (2010) where OCD participants were less accurate compared to controls in a similar task. This deficit was independent of the feedback direction, i.e. negative feedback block, positive feedback block and a mixed feedback. However, their task was different because it had a within participant design instead of a between participant design, it used less trials (32-36 trials), and feedback points were given instead of monetary feedback. Finally, Starcke et al. (2010) investigated OCD patients instead of subclinical checker participants. Hence, it might be possible that Starcke et al. (2010) results are due to these experimental differences, esp. sample differences.

In addition to the accuracy analysis in the random sequence block, the usage of exploration/exploitation strategies was evaluated. Findings showed no exploration/exploitation differences between the subclinical checker group and the non-checker group in both experiments, and also no exploration/exploitation differences between experiments. Actually, this indicates that checkers and non-checkers are similarly influenced by errors and winnings in this task and that their behaviour does not depend on the feedback direction used in the RBL task. In other words, checkers who were supposed to be more influenced by negative monetary feedback seem to be similarly affected by it in comparison to positive monetary feedback in a random task.

An explanation for these results is that OCD deficits in RBLTs are supposedly not related with feedback direction but with other factor commonly present in RBLT, as feedback magnitude or feedback delay (see Chapter 6 for more information). In this way, positive and negative direction would similarly affect checkers and non-checkers. Moreover, it is also possible that subclinical checkers might not present enough obsessive compulsive symptoms to be affected by the feedback direction of the task and that clinical individuals need to be tested to find performance related differences. Finally, it is possible that checkers might only be affected by symptom-related negative feedback but not general aversive feedback in a RBLT. This will be furtherly discussed in Chapter 7.
Regarding correlations between the performance in the random sequence block and the questionnaires, results revealed that (1) checkers with higher symptom severity had a better performance in block 1 in the positive neutral experiment, (2) that anxiety was related to enhanced lose-shift behaviour in the positive-neutral experiment, and (3) that financial risk-taking is related to more shifts after winning in the negative neutral experiment, which is related to higher auto- and cross-correlation differences in lag 1 because risk-takers seem to be more influenced by previous winnings.

If these correlations were common between the negative-neutral and positive-neutral experiment, they could be used to explain the results, especially because there was no difference between experiments for checkers in block 1. However, while binary decision making variables were similar between experiments, some correlations were only found for the positive-neutral experiment in block 1, while others for the negative-neutral experiment. In this way, it is highly possible that these results might be more related with specific sample characteristics of the two experiments than with something meaningful regarding this investigation. Nevertheless, these correlations were also recorded in Chapters 6 and 7 to investigate if they might be able to explain group or experimental differences in one of the studies which might be related to a specific task factor, such as feedback-magnitude in Chapter 6 and symptom-related feedback in Chapter 7. They will be discussed together with results from Chapters 6 and 7 in the general discussion presented in Chapter 8.

The second aim of the study was to extend the investigation of probabilistic reward-based learning tasks from random to pattern sequences. The findings from the pattern sequence block 2 were rather interesting. Checkers seemed to be able to learn the pattern sequence of block 2 better compared to non-checkers in the negative-neutral experiment. There were no group differences for the positive-neutral experiment. These differences were present from the second sub-block onwards (after 32 trials). In accordance, the auto-and cross-correlation analysis also showed that checkers seemed to present a better performance compared to non-checkers in the negative-neutral experiment, and that they were able to learn the pattern better in the negative-neutral experiment compared to the positive-neutral experiment. Moreover, the strategy questionnaire revealed that checkers reported seeing more patterns in the negative-neutral experiment.

These results seem to confirm one of this study hypothesis that checkers would present a better performance in the negative-neutral experiment in block 2 than non-checkers but not in the positive-neutral experiment, possibly because checkers might be more biased
towards exploitation in a negative feedback situation where the negative feedback would influence checkers to decrease risky choices; which is a more optimal strategy in the task. More specifically, the bias towards exploitation in checkers might decrease the probability of checkers to be influenced by the disruptions presented in the pattern sequence which would increase their performance. In contrast, non-checkers might be more prone to shift options and test new hypothesis after a pattern disruption which would reflect a stronger exploration bias towards new hypotheses, decreasing performance. This explanation hypothesis could be tested by designing a new study where checkers and non-checkers would be exposed to different experiments containing this same pattern sequence with distinct probabilities of disruption. If this explanation is correct, checkers and non-checkers would equally perform in a sequence where disruptions are inexistent but differences between groups would be increased after adding more and more disruptions to the pattern. Definitely, it is important to consider that too many disruptions would turn the pattern impossible to learn, so there might be a limit of adding disruptions in the pattern, where both groups would be unable to learn the optimal strategy of the task. In this case, no differences would be found between groups either. Clearly, if checkers are supposedly more affected by negative feedback than positive feedback, these results would be only true in experiments with negative feedback.

Furthermore, the correlation findings for the negative-neutral experiment (study 3) showed that higher levels of checking and intolerance of uncertainty might be related to reduced pattern detection, as reflected in the cross-correlations. This finding might point out that high checking symptom levels could reduce pattern learning in a RBLT. This could be tested with clinical checkers with high checking symptom levels. Indeed, a potential hypothesis would be that checking symptoms increase the bias towards exploitation which is initially beneficial for pattern learning. However, the bias becomes costly with increasing symptom severity because, eventually, patients would not be able to learn the pattern sequence in the first place, as for the initial pattern learning, a minimum of exploration behaviour is required. In addition, intolerance of uncertainty levels might have the same effect on performance because participants with high intolerance of uncertainty levels would choose the safest strategy, in other words they would be biased towards exploitation. A very high exploitation bias would hinder participants to learn the optimal strategy of the task.

Regarding block 3 results, no group or experimental accuracy and cross-correlation differences were found in this block where target option maximisation and pattern search strategies were combined. Note, there was a significant correlation between OCI-R total and
accuracy in block 3 which will be further discussed in Chapter 8. The behavioural findings in block 3 are similar to findings in block 1. They indicate that feedback direction might not be an explanation for the RBL deficits in OCD when target option probabilities are manipulated, independent of the absence or presence of a pattern in the event sequence.

In contrast, results from the pattern block with a chance level target option probability (block 2) seem to confirm the hypothesis that checkers would be more biased towards exploitation in a negative-neutral experiment. It is possible that this bias towards exploitation is also happening for blocks 1 and 3 in checkers. However, no effects of group and experiments were found for blocks 1 and 3. One possible explanation for feedback direction effects in block 2, but not in blocks 1 and 3, is that non-checkers might be more biased towards exploitation in blocks 1 and 3 than in block 2, so differences between groups were more evident for block 2. Indeed, this block presents an optimal strategy that is easier to learn for participants than the optimal strategy of block 1 and 3 because its maximum accuracy is higher than block 1 (0.82 for block 2 and 0.72 for block 1) and it only presents one optimal strategy (pattern) and not two, as block 3 (more target in one option and pattern). In this way, it is possible that non-checkers might be more biased towards exploitation in blocks 1 and 3 because these blocks are more complex to learn, so it would be less probable to find differences between groups for these blocks.

For now, it could be argued that checkers might present a deficit in RBLTs that use negative feedback directions when learning is solely based on the presence and usage of pattern sequences (non-random conditional probabilities) with specific values of maximum accuracy. However, it could also be argued that the effect of the negative feedback direction might be reflecting a deficit regarding a specific error magnitude in the negative-neutral experiment. Indeed, from this study, it is not possible to conclude that checkers present a deficit related with a general negative feedback direction. It is possible that in others negative feedback magnitudes, checkers might not present these deficits, for instance. In the other hand, if checkers supposedly do present biased emotional responses towards aversive emotions in general, checkers would present a general bias towards exploitation irrespectively of the magnitude of the negative task and would not present any deficit, regarding tasks within a positive feedback direction, even if the magnitude of the error of these positive tasks were manipulated.

Note that, while the positive-neutral experiment presents rewards, it also offers an absence of reward after an error that can be interpreted as an aversive emotional stimuli. In
this way, it is possible that the manipulation of the error magnitude associated with the absence of reward in the positive feedback direction could also influence checkers to be more biased towards exploitation than non-checkers as it has happened in block 2 for the negative-neutral experiment. Indeed, this investigation about the influence of distinct values of negative feedback magnitude in checkers will be pursued in Chapter 6, where the influence of the feedback magnitude manipulation within a positive feedback direction and the influence of the feedback magnitude manipulation within a negative feedback direction will be separately examined. Note that positive and negative experiments will not be investigated together, as previous literature has shown that negative feedback associated with an absence of a reward and negative feedback associated with the presence of a penalty are related with distinct physiological systems (Schultz, 2010).

5.4.1. Limitations and future directions

There are several potential limitations of this first study. Firstly, the most obvious limitation is a small sample size and potential sampling errors in a between-participant design. Hence, it would be important to replicate the findings from this study, considering a larger sample size of subclinical checkers. Secondly, the use of a subclinical checker sample with a lower symptom severity could be another reason for negative, non-significant, findings. Therefore, it would be useful to replicate this study with a clinical sample, taking into consideration the fact that a clinical sample might be less able to learn the task, as shown by Starcke et al. (2010) in their study. Of course, a clinical sample has other limitations such as medication issues, higher levels of comorbidity, and therapeutic treatment influences.

Thirdly, no direct physiological measures of emotional sensitivity to negative and positive feedback direction was taken, e.g. skin conductance responses (SCR) and heart rates. Actually, it would be interesting to replicate this study using SCR and heart rate measurements, so it would be possible to correlate accuracy and cross-correlation values in block 2 with these physiological measures.

5.5. Conclusions

In this study, checkers and non-checkers performed a binary decision making task with either positive vs. absent feedback or negative vs absent feedback. The findings indicated that checkers might be as able to learn probabilistic event sequences with and without patterns as non-checkers and they supposedly were actually better learners when only
pattern information was provided in the negative-neutral experiment. In this block, checkers seemed to be more biased towards exploitation. Interestingly, correlations indicated that checkers with higher symptom severity and intolerance of uncertainty got worse in the pattern block. This correlational finding needs to be studied in more detail in future experiments before making any conclusions.

As these group and experimental differences were only found for block 2 but not for blocks 1 and 3, as predicted, it might be possible that the exploitation bias in checkers and non-checkers depend on feedback direction but also others variables, as maximum accuracy values and the sole presence of a unique optimal strategy, as patterns. Investigations of these explanations could be considered for future studies. However, the investigation of the effect of feedback direction in the decision making of checkers will be pursued in Chapter 6, now considering the investigation of how the manipulation of the feedback magnitudes can influence checkers.

In Chapter 6, we investigated whether differences regarding block 2 are enhanced with higher feedback magnitude values for both feedback directions. To investigate this further, new experiments were designed to examine how feedback magnitude manipulation can affect RBL in checkers in comparison to non-checkers. As the value comparison of a feedback magnitude for the presence of a penalty (negative experiments) with the absence of a reward (positive experiments) is complicated because they represent distinct variables, two studies were developed to investigate the effect of feedback magnitude manipulation in checkers, one considering the manipulation of error feedback magnitudes associated with a positive feedback (absence of rewards) and the other considering the manipulation of negative feedback magnitudes (presence of penalties). These studies will be furtherly discussed in Chapter 6.
Chapter 6 Reward based-learning in subclinical checkers: The influence of feedback magnitude

6.1. Introduction

OCD population seems to present a specific decision making deficit regarding reward-based learning (Starcke et al., 2010). However, findings were contradictory as studies report deficits and sometimes they do not (for example, see Starcke et al. (2010) versus Gründler et al. (2009), respectively).

These ambiguous findings could be explained by several reasons, regarding 1) the sample used in each study or 2) the tasks employed. In fact, some studies employed samples where the use of medication is not controlled, while others do not consider the presence or absence of common comorbidities associated with OCD as depression and anxiety. Finally, some studies do not differentiate between distinct OCD subtypes, and, the tasks used in these studies are generally different. Regarding the tasks employed, three limitations could explain contradictory findings.

One limitation of the previous studies regards the fact that these tasks often manipulate several factors at the same time, as feedback direction (positive vs. negative vs. mixed feedback during learning), feedback magnitude (e.g. amount of positive vs. negative feedback), feedback gain (difference in feedback between options), or feedback delay (immediate vs. delayed reward), so it is not possible to know which of these factors are in deficit in OCD.

A second limitation is related with the restrict use of accuracy as an analysis to investigate possible RBL deficits in OCD, not even considering an analysis of learning over time (learning curves) or the use of previous results to make future predictions by participants. Reward-based learning involves the use of prediction errors that are a comparison of actual results with previous unexpected results during decision making (Maia, 2009). Unfortunately, analysis of the use of previous trials during decision making in a RBLT are usually absent in investigations of RBL deficits in OCD but these analysis could indicate deficits in prediction errors in populations that seem to present RBL deficits, as OCD. Actually, this could be done by measuring the probability a participant shifts options after losing or winning in the immediate previous trial.

Finally, a third limitation concerns the absence of the investigation of how OCD is affected by the use of one or more previous trials to make predictions, a phenomenon called
sequential effects (Wilder at al., 2009). According to the authors, RBL do not only involves the use of the immediate previous trial to make predictions but the use of other previous trials (sequential effects), as a way to discover possible patterns in the sequence (Vulkan, 2000). Indeed, deficits in the use of sequential effects could influence performance in OCD, as this effect might be also present in random sequence tasks (Vulkan, 2000; Unturbe & Corominas, 2007). In fact, it is possible to investigate how participants are affected by one or more previous trials, measuring the probability of correlation between one actual response to a previous response (cross-correlation). However, more than examining the use of previous trials by participants in a random sequence, it is interesting to investigate these sequential effects with event sequences that have hidden patterns to find out how people with OCD or subclinical OCD symptoms deal with these pattern sequences.

In this way, considering these limitations from previous studies, a new binary decision making task was designed. In Chapter 5, a study tried to investigate RBL deficits in subclinical checkers related to feedback direction (study 2 as positive-neutral vs. study 3 as negative-neutral). Random and pattern sequences were considered and analysis of the use of previous trials were employed (probability of shifting and auto- cross-correlations). Based on evidence that OCD is more sensitive to aversive stimuli (Simon et al., 2010), it was expected that subclinical checkers would present a specific deficit in the negative condition but not in the positive condition. Indeed, it was expected that subclinical checkers would present a risk-aversive behaviour in the negative experiment, as a way to avoid new negative feedback, what would increase exploitation of the optimal strategy in subclinical checkers, increasing performance.

Findings from Chapter 5 showed that checkers do not present a specific deficit in a RBL task with random sequences, irrespectively of the feedback direction. However, the picture was different when analysing the pattern sequence block, where subclinical checkers learned the pattern sequence in the negative-neutral experiment (study 3) more efficiently than non-checkers. These results partially confirmed the hypothesis that subclinical checkers would present a bias towards exploitation in the negative direction but it is still not possible to know if this deficit is more related with the feedback direction of the task or with the specific feedback magnitude of the negative-neutral experiment. In other words, it is possible that subclinical checkers do present a deficit regarding tasks within a negative feedback direction but it is also possible that the feedback magnitude of the task is responsible for this deficit, so
a higher negative feedback magnitude associated with the error of a positive task (absence of reward), for instance, would also affect subclinical checkers.

If this is the case, higher negative feedback magnitudes than the ones presented in the negative-neutral experiment (study 3) might affect the bias towards exploitation in subclinical checkers. Additionally, higher negative feedback magnitudes than the one presented in the positive-neutral experiment could also bias subclinical checkers towards exploitation. If feedback magnitude is not related with results found in Chapter 5 but only feedback direction, a bias towards exploitation will be only found in the experiments within the negative direction, irrespectively of the manipulation of the error magnitude of these tasks. Moreover, there might not be any deficit associated with the positive tasks.

To examine these hypothesis, it was first important to consider that while it is highly intuitive that an error in the positive-neutral experiment presents a lower magnitude than an error in the negative-neutral experiment, there is no guarantee that the absence of a reward in the positive-neutral experiment is felt as less negative than the presence of a penalty in the negative-neutral experiment. In others words, it is not possible to compare the positive experiments with the negative ones in terms of error magnitude. This is the case because errors associated with the positive experiments are related with an absence of feedback and errors in the negative experiments are related with the presence of a penalty. Indeed, there is evidence that there are two separate physiological systems for prediction errors related with the absence of rewards and prediction errors related with the presence of penalties (Fiorillo, 2013; Schultz, 2010). For these reasons, it is important to separately investigate how checker decision making is affected when distinct error magnitudes are manipulated, either in a positive direction or in a negative direction.

In this way, the first aim of this study was to investigate RBL in subclinical checkers while exposed to distinct amounts of error magnitudes in two positive experiments. To do so, a positive-positive experiment (study 4) was designed, where participants received money and saw a positive picture (two rewards) after finding the target and participants did not receive money and did not see a positive picture after an error (double absence of reward). This experiment was compared with the positive-neutral experiment (study 2), so two magnitudes of rewards and, consequently, two magnitudes of absence of rewards were compared. Note that as in the positive-neutral experiment (study 2), participants did not lose anything after a mistake, so mistakes were always represented by an absence of something.
Our second aim was to investigate how checkers were affected by distinct amounts of penalties after an error, so a negative-negative experiment (study 5) was designed, where participants lost money and saw a negative picture (two penalties) after a mistake and participants did not lose money and saw a neutral picture after finding the target (two absence of penalties) (study 5). This experiment was compared with the negative-neutral experiment (study 3), so two magnitudes of penalties were compared and consequently, two magnitudes of absence of penalty. Note that the reward was always represented by the absence of penalties. Emotional pictures were chosen for the feedback magnitude enhancement because this allowed us to manipulate feedback magnitudes by using symptom-related pictures in future experiments like in the study presented in Chapter 7.

Finally, the third aim of this study was to investigate changes in strategy, in addition to accuracy, regarding first- and second order sequential effects in the decision making of OCD and how these strategies were particularly affected by the feedback magnitude manipulation in the positive and in the negative direction. Random-sequence and pattern sequence blocks were used to investigate these effects. Indeed, these sequences were exactly the same as in Chapter 5, so each experiment was divided in three blocks, each one presenting a distinct probabilistic environment. In block 1, there were no sequential effects (no pattern) but a higher probability of winnings in one of the target options. In block 2, a pattern sequence was presented, e.g. LLRR, but additional noise was added to make the detection of the pattern more difficult. Both target options had an equal probability of winning and performance could only be improved by implicitly learning the pattern in the event sequence. In block 3, there was a pattern and differential target option probabilities which were the same as in block 1.

Based on previous results and on previous literature associating OCD with a higher sensitivity to negative feedback, it is possible that OCD individuals will present a bias towards exploitation when error magnitudes associated with the task are too high (i.e., more negative). Actually, results from Chapter 5 showed that subclinical checkers seem to only present an altered performance in the negative-neutral experiment, so it is possible that this deficit depends on the feedback direction of the task or on the feedback magnitude of the task. In this way, it is possible to hypothesize that 1) if feedback direction is affecting checker RBL and this is independent from the feedback magnitude when within the same feedback direction, subclinical checkers will not present any difference in accuracy, probability of shifting and cross-correlation between positive experiments (study 2 as positive-neutral vs.
study 4 as positive-positive) and between negative experiments (study 3 as negative-neutral vs. study 5 as negative-negative). However, 2) if feedback magnitude and not feedback direction is affecting checker RBL, differences between experiments in the positive or in the negative studies might be found.

If the second hypothesis is correct and feedback magnitude is affecting checker performance, higher error magnitudes will affect checker performance because they seem to be more sensitive to aversive stimuli than to positive emotional stimuli (Simon et al., 2010). Previous results from Chapter 5 have shown that checker performance was not affected by the error magnitude of the positive-neutral experiment, so if the error magnitude associated with the positive-positive experiment is too aversive for checkers, differences between experiments will be found for accuracy, probability of shifting and cross-correlation, where checkers will present a bias towards exploitation in the positive-positive experiment. However, if the error magnitude associated with the positive-positive experiment is not aversive enough, there will not be any experimental difference for checkers in the positive study.

In addition, for the comparison between the negative experiments (study 3 vs. study 5), if the feedback magnitude of the negative-neutral experiment is already affecting checker performance, it will be expected that increasing the feedback magnitude will increase even more the bias towards exploitation in the negative-negative experiment (study 5) and differences between experiments for checkers will be found. However, it is also possible that increasing the feedback magnitude from the negative-neutral (study 3) to the negative-negative experiment (study 5) will not augment the bias towards exploitation, so no differences will be found between experiments in the negative study.

Note that this bias towards exploitation in the negative-neutral experiment was only the case in block 2, so results in this study might be also dependent on the maximum accuracy of the task. In this way, it is possible that these hypothesis will be only the case for block 2 and blocks 1 and 3 will not present any distinction between experiments and groups.

As the comparison between the positive experiments and the comparison between the negative experiments were investigated separately, results for the positive experiments will be shown first as a positive study (study 2 vs. study 4) and the negative experiment results will follow as a negative study (study 3 vs. study 5). Results will be discussed together.
6.2. Methods

6.2.1. Overview of the positive and negative study design

The study design was exactly the same as the one presented in Chapter 5. Again, students completed an online screening questionnaire and potential participants were screened based on the inclusion and exclusion criteria (see Chapter 2). Selected participants were randomly assigned to perform the positive-positive (study 4) or the negative-negative experiment (study 5) in the laboratory. The laboratory study did consist of three parts, the picture validation task, a binary decision making task (either the positive-positive experiment or the negative-negative experiment), and the completion of the laboratory-based questionnaires.

Results from the positive-positive were compared with the positive-neutral experiment (study 2 vs. study 4) and results from the negative-negative were compared with the negative-neutral experiment (study 3 vs. study 5).

6.2.2. Participant selection based on the online screening questionnaire

The Online screening questionnaire consisted of some demographic questions about age, gender, play videogame and play for money, past or present diagnosis of depression, anxiety, panic disorder and post-traumatic stress disorder, the Obsessive Compulsive Inventory revised (OCI-R) and the Depression, anxiety and stress scales (DASS) questionnaires. Methods about the online questionnaire can be seen in detail in Chapter 2.

352 people signed in for the study. From this amount, 154 did not complete the online questionnaire. In this way, 194 online questionnaires were further analysed and from this 94 participants were excluded on the basis of the study exclusion criteria and 100 were invited for the lab-based study. Of these invited participants, 72 data from participants were analysed. 24 participant’s data were analysed in the negative-negative experiment, 24 of them in the positive-positive and 24 of them in the symptom-related experiment (shown in Chapter 7). Symptom-related results will be shown in Chapter 7 but participants were previously recruited at the same time. In each experiment, 12 participants were checkers and 12 were non-checkers. Overall, participants were excluded in part 1 for a number of reasons related to the study exclusion criteria (see table 6.1). Participants were also excluded after selection and completion of part 2. Reasons can also be found at table 6.1.
Table 6.1.
*Reasons and number of participants excluded and included in the positive, negative and symptom-related study (Chapter 7)*

<table>
<thead>
<tr>
<th>Exclusion reason</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total participants</td>
<td>352</td>
</tr>
<tr>
<td>Incomplete questionnaires</td>
<td>154</td>
</tr>
<tr>
<td>Questionnaire completed twice</td>
<td>4</td>
</tr>
<tr>
<td>Questionnaire completed</td>
<td>194</td>
</tr>
<tr>
<td>Age</td>
<td>9</td>
</tr>
<tr>
<td>Diagnosed anxiety disorder (controls)</td>
<td>3</td>
</tr>
<tr>
<td>Diagnosed depression</td>
<td>17</td>
</tr>
<tr>
<td>Diagnosed depression &amp; anxiety</td>
<td>8</td>
</tr>
<tr>
<td>Diagnosed depression &amp; panic disorder</td>
<td>3</td>
</tr>
<tr>
<td>Diagnosed depression, anxiety &amp; PTSD</td>
<td>1</td>
</tr>
<tr>
<td>Diagnosed depression, anxiety, panic disorder &amp; PTSD</td>
<td>1</td>
</tr>
<tr>
<td>DASS score</td>
<td>14</td>
</tr>
<tr>
<td>Neither checker, nor non-checker</td>
<td>38</td>
</tr>
<tr>
<td>Remaining participants after inclusion / exclusion</td>
<td>100</td>
</tr>
<tr>
<td>No response to invitation to lab-based study</td>
<td>22</td>
</tr>
<tr>
<td>Recorded in the lab-based study</td>
<td>78</td>
</tr>
<tr>
<td>Exclusion from the lab-based study</td>
<td>6</td>
</tr>
<tr>
<td><strong>Complete results</strong></td>
<td><strong>72</strong></td>
</tr>
</tbody>
</table>

6.2.3. **Participants in the laboratory-based study: Group sample characteristics**

For tests of normality, z-scores for skewness and kurtosis were calculated for age, and the DASS and OCI-R scores. Gender, videogame experience, habit to play for money were investigated as categorical variables.

DASS depression for the negative-negative experiment was not a normally distributed variable. For the OCI-R questionnaire, OCI-R washing and OCI-R obsessing in the positive-positive experiment and OCI-R checker, OCI-R washing, OCI-R ordering and OCI-R counting in the negative-negative experiment, variables were not normally distributed.

Online questionnaire variables were compared across groups and experiments. For most variables, this comparison was conducted by using between-participant ANOVAs with
the factors Group (checker vs. non-checker) and experiment (positive-neutral vs. positive-positive OR negative-neutral vs. negative-negative). Gender, videogame playing and gambling with money were the exceptions because these were categorical variables. Here, Chi-Square tests or Fisher’s exact tests were used to explore sample differences between groups or experiments. The findings from this analysis are presented in Table 6.2.

For the Positive study (positive-neutral vs. positive-positive), for both the OCI-R and the DASS scores, ANOVAs did reveal significant main effects of Group (Table 6.2). Basically, checkers always had higher scores than non-checkers in all OCI-R and DASS subscales, independent of the experiments (Table 6.2). For age, however, non-checkers (22.43 ± 0.79) were slightly older than checkers (19.96 ± 0.33), main effect of Group: F(1, 42) = 7.88, p=.008). However, this difference is unlikely to affect performance, as previous studies showed young adults from 20-25 years old present similar decision making strategies (Victorino et al., unpublished). In addition, a significant interaction between group and experiment was found for the OCI-R washing rating scale (Table 6.2). More specifically, checkers also had enhanced washing symptoms compared to non-checkers in the positive-positive experiment (Mean (Checker): 4.08±0.66; Mean (Non-checker): 0.5±0.15; U = 6, z = -3.91, p_b =.002, r = .83) but there was no group difference for the positive-neutral experiment (Mean (Checker): 2.36±0.62; Mean (Non-checker): 1.18±0.26, t(13.47) = 1.75, p_b = .10, r = .83).

For the Negative Study (negative-neutral vs. negative-negative), for both the OCI-R scores and the DASS scores, ANOVAs did reveal significant main effects of Group (Table 6.3). Basically, checkers always had higher scores than non-checkers for the OCI-R total and all OCI-R subscales, except washing, and for all DASS subscales. (Table 6.3). There were no significant group differences for gender, play videogame and play for money. Differently from the positive study, there was no main effect of group for age (Table 6.3). Note, for all assessed scales and descriptive variables, no significant differences between the negative-neutral and negative-negative experiment were found, as reflected in the absent main effects of Experiment. There were also no significant interactions between Group and Experiment (Table 6.3).

These findings were similar to Chapter 5 and showed that the participant selection process worked well for these experiments and that groups can be compared across experiments. Sampling differences between experiments cannot explain potential effects in the behavioural measures.
Table 6.2.
Mean, standard error and statistics for checkers and non-checkers in the positive-neutral and positive-positive experiments, considering all variables analysed for the investigation of sample characteristics. $F$ = female, $M$ = male, $Pv$ = videogame player, $NPv$ = no videogame player, $Pm$ = money player, $NPm$ = no money player; $ME$ = main effect; $IA$ = interaction

<table>
<thead>
<tr>
<th>Variables</th>
<th>Positive-neutral</th>
<th>Positive-positive</th>
<th>Chi-square or Fisher’s exact test</th>
<th>Checkers (between experiments)</th>
<th>Non-checkers (between experiments)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Checkers</td>
<td>Non-checker</td>
<td>Checkers</td>
<td>Non-checker</td>
<td>Positive-neutral (between groups)</td>
</tr>
<tr>
<td>Demographic questionnaire (Categorical variables)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>5 F, 6 M</td>
<td>9 F, 2 M</td>
<td>7 F, 5 M</td>
<td>8 F, 4 M</td>
<td>$p = .18$</td>
</tr>
<tr>
<td>Play videogame</td>
<td>5 Pv, 6 NPv</td>
<td>6 Pv, 5 NPv</td>
<td>7 Pv, 5 NPv</td>
<td>10 Pv, 2 NPv</td>
<td>$p = .67$</td>
</tr>
<tr>
<td>Play for money</td>
<td>2 Pm, 9 NPm</td>
<td>2 Pm, 9 NPm</td>
<td>1 Pm, 11 NPm</td>
<td>0 Pm, 12 NPm</td>
<td>$p = 1$</td>
</tr>
<tr>
<td>Demographic questionnaire (Continuous variables)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>20.36± 0.60</td>
<td>22.00 ± 1.15</td>
<td>19.58± 0.31</td>
<td>22.83± 1.11</td>
<td>$F(1, 42) = 0.001$, $p=.97$</td>
</tr>
<tr>
<td>OCI-R scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

162
<table>
<thead>
<tr>
<th>OCI-R behavior</th>
<th>Mean1 ± SD1</th>
<th>Mean2 ± SD2</th>
<th>Mean3 ± SD3</th>
<th>Mean4 ± SD4</th>
<th>F(1, 42)</th>
<th>p</th>
<th>η^2</th>
<th>p_group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Checking</td>
<td>7.36±0.59</td>
<td>0.81±0.22</td>
<td>7.33±0.72</td>
<td>0.58±0.22</td>
<td>0.07</td>
<td>.79</td>
<td></td>
<td>.80</td>
</tr>
<tr>
<td>Washing</td>
<td>2.36±0.62</td>
<td>1.18±0.26</td>
<td>4.08±0.66</td>
<td>0.5±0.15</td>
<td>1.15</td>
<td>.28</td>
<td></td>
<td>.36</td>
</tr>
<tr>
<td>Hoarding</td>
<td>4.81±0.68</td>
<td>2.36±0.57</td>
<td>5.58±0.73</td>
<td>1.58±0.54</td>
<td>0</td>
<td>.99</td>
<td></td>
<td>.37</td>
</tr>
<tr>
<td>Ordering</td>
<td>6.27±1</td>
<td>2.81±0.60</td>
<td>6.75±0.68</td>
<td>2.33±0.63</td>
<td>0</td>
<td>.99</td>
<td></td>
<td>.39</td>
</tr>
<tr>
<td>Counting</td>
<td>2.63±0.80</td>
<td>0.27±0.19</td>
<td>3.83±0.85</td>
<td>0.42±0.14</td>
<td>1.25</td>
<td>.27</td>
<td></td>
<td>.35</td>
</tr>
<tr>
<td>Obsessing</td>
<td>3.63±0.52</td>
<td>1.00±0.30</td>
<td>3.83±0.72</td>
<td>0.42±0.19</td>
<td>0.15</td>
<td>.69</td>
<td></td>
<td>.47</td>
</tr>
<tr>
<td>TOTAL</td>
<td>27.09±2.30</td>
<td>8.45±1.32</td>
<td>31.42±2.28</td>
<td>5.83±1.22</td>
<td>0.21</td>
<td>.64</td>
<td></td>
<td>.77</td>
</tr>
</tbody>
</table>

**DASS scores**

<table>
<thead>
<tr>
<th>DASS Depression</th>
<th>Mean1 ± SD1</th>
<th>Mean2 ± SD2</th>
<th>Mean3 ± SD3</th>
<th>Mean4 ± SD4</th>
<th>F(1, 42)</th>
<th>p</th>
<th>η^2</th>
<th>p_group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>5.72±1.33</td>
<td>2.63±0.85</td>
<td>2.83±0.73</td>
<td>2.00±0.85</td>
<td>3.38</td>
<td>.07</td>
<td></td>
<td>.09</td>
</tr>
<tr>
<td>Anxiety</td>
<td>6.27±1.28</td>
<td>1.09±0.34</td>
<td>5.5±1.35</td>
<td>1.67±0.67</td>
<td>0.01</td>
<td>.92</td>
<td></td>
<td>.31</td>
</tr>
<tr>
<td>Stress</td>
<td>9.18±1.67</td>
<td>3.54±0.74</td>
<td>7.92±1.40</td>
<td>4.17±1.50</td>
<td>0.05</td>
<td>.81</td>
<td></td>
<td>.21</td>
</tr>
</tbody>
</table>
Table 6.3.
Mean, standard error and statistics for checkers and non-checkers in the negative-neutral and negative-negative experiments, considering all variables analysed for the investigation of sample characteristics. F = female, M = male, Pv = videogame player, NPv = no videogame player, Pm = money player, NPm = no money player; ME = main effect; IA = interaction

<table>
<thead>
<tr>
<th>Variables</th>
<th>Negative-neutral</th>
<th>Negative-negative</th>
<th>Chi-square or Fisher’s exact test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Checkers</td>
<td>Non-checker</td>
<td>Checkers</td>
</tr>
<tr>
<td>Demographic questionnaire (Categorical variables)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>10 F, 2 M</td>
<td>9 F, 3 M</td>
<td>10 F, 2 M</td>
</tr>
<tr>
<td>Play videogame</td>
<td>6 Pv, 6 NPv</td>
<td>9 Pv, 3 NPv</td>
<td>7 Pv, 5 NPv</td>
</tr>
<tr>
<td>Play for money</td>
<td>1 Pm, 11 NPm</td>
<td>0 Pm, 12 NPm</td>
<td>0 Pm, 12 NPm</td>
</tr>
<tr>
<td>Demographic questionnaire (continuous variables)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>20.25± 0.49</td>
<td>22.91± 0.6</td>
<td>21.17± 0.58</td>
</tr>
</tbody>
</table>

OCI-R scores
| OCI-R Checking | 7.25±0.68 | 0.91±0.25 | 5.92±0.41 | 1±0.27 | F(1, 44) = 1.98, p = .16 | F(1, 44) = 160.50, p < .001** | F(1, 44) = 2.54, p = .11 |
| OCI-R Washing | 2.66±0.72 | 1.08±0.68 | 1.5±0.57 | 0.92±0.39 | F(1, 44) = 1.22, p = .27 | F(1, 44) = 3.23, p = .08 | F(1, 44) = 1.07, p = .30 |
| OCI-R Hoarding | 4.50±0.67 | 1.50±0.43 | 6.17±0.94 | 1.75±0.42 | F(1, 44) = 1.97, p = .16 | F(1, 44) = 29.52, p < .001** | F(1, 44) = 1.07, p = .30 |
| OCI-R Ordering | 5.41±0.60 | 2.58±0.63 | 5.75±1.01 | 1.08±0.41 | F(1, 44) = 0.68, p = .41 | F(1, 44) = 28.44, p < .001** | F(1, 44) = 1.7, p = .19 |
| OCI-R Counting | 3.91±0.91 | 0.33±0.14 | 3.08±0.71 | 0.33±0.25 | F(1, 44) = 0.48, p = .49 | F(1, 44) = 27.98, p < .001** | F(1, 44) = 0.48, p = .49 |
| OCI-R Obsessing | 3.91±0.89 | 1.16±0.42 | 4.33±0.79 | 1.33±0.33 | F(1, 44) = 0.09, p = .76 | F(1, 44) = 16.97, p < .001** | F(1, 44) = 0.03, p = .85 |
| OCI-R Total | 27.66±3.16 | 7.58±1.25 | 26.67±2.63 | 6.33±1.32 | F(1, 44) = 0.24, p = .62 | F(1, 44) = 78.26, p < .001** | F(1, 44) = 0.03, p = .95 |

| DASS Scores | |
| DASS Depression | 6.08±1.29 | 2.25±1.05 | 2.67±0.82 | 1.92±0.57 | F(1, 44) = 3.92, p = .054 | F(1, 44) = 5.86, p = .02* | F(1, 44) = 2.65, p = .11 |
| DASS Anxiety | 6.08±1.29 | 2.41±0.74 | 3.92±0.97 | 1.83±0.60 | F(1, 44) = 2.14, p = .15 | F(1, 44) = 9.35, p = .004** | F(1, 44) = 0.70, p = .40 |
| DASS Stress | 11.16±1.67 | 3.50±1.06 | 9.33±2.25 | 3.92±1.10 | F(1, 44) = 0.19, p = .76 | F(1, 44) = 16.75, p < .001** | F(1, 44) = 0.49, p = .48 |

η² group = .78

η² group = .40

η² group = .39

η² group = .38

η² group = .27

η² group = .64

η² group = .11

η² group = .17

η² group = .27
6.2.4. Laboratory-based Study

6.2.4.1. Laboratory-based study procedure

These experiments were conducted at the University of Surrey. Upon arrival at the laboratory, participants received the Participant Information Sheet (PIS), containing information about the experiment. There were two possible groups participants could be randomly assigned to, the positive-positive and the negative-negative experiment. Depending on the group, participants received either a PIS regarding the positive-positive experiment or a PIS regarding the negative-negative experiment (Appendix 2.4 for an example of PIS). After reading it, participants did read and sign the Consent form (Appendix 2.5). The total lab-based study took approximately two hours. It consisted of the standard Picture Validation task, a binary decision making task, and five laboratory-based questionnaires which were the Domain Specific Risk-Taking Scale (DOSPERT); Intolerance of uncertainty questionnaire (IUS); Impulsive Behaviour Scale (UPPS), the South-Oaks Gambling Screen, and a strategy questionnaires (see Chapter 2 for more details).

The Picture Validation task was the same as described in Chapter 2. The binary decision making tasks were similar to the one described in Chapter 2 but they differed in their feedback type. The tasks were designed to investigate the effect of feedback magnitude within the positive direction (positive study as study 2 vs. study 4) and negative direction (negative study as study 3 vs. study 5) on RBL. The specific feedback magnitudes for the two experiments were positive monetary feedback and positive pictures for the positive-positive experiment and negative monetary feedback and negative pictures for the negative-negative experiment.

6.2.4.2. Data analysis for the laboratory-based study

Data analysis for the positive study and negative study was exactly the same as the one employed at Chapter 5.2.4.2. As a short reminder, for the picture validation task, data analysis considered measures of valence, arousal and comfort. For each measure, a three-way 4 x 2 x 2 mixed ANOVA with the within-participant factor Picture Type (positive, neutral, negative and checking-related pictures) and the between-participant factor Group (checkers vs. non-checkers) and Experiment was conducted for each study. Data analysis of the binary decision making task investigated accuracy, sub-block accuracy, probability of shifting after winning and after losing in the previous trial and cross-correlations between the stimulus and
the response sequence for the five previous trials (lags). The exact type of statistical test employed for each one of these analysis can be seen at table 6.4. For all measures, all three blocks were analysed separately because of their maximum accuracy and structural differences. The exception is the exploration / exploitation (win shift / lose shift) measure where only block 1 was analysed.

Table 6.4

*Statistical tests employed for each analysis conducted for the binary decision making task in the positive and negative study*

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Within-participant factor</th>
<th>Between-participant factor</th>
<th>Type of test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>None</td>
<td>Group</td>
<td>Two-way 2 x 2 ANOVA</td>
</tr>
<tr>
<td>Sub-block accuracy</td>
<td>5 Sub-blocks</td>
<td>Checkers vs. non-checkers</td>
<td>Three-way 5 x 2 x 2 mixed-ANOVA</td>
</tr>
<tr>
<td>Win-shift</td>
<td>None</td>
<td>Positive study (Positive-neutral vs. positive-positive)</td>
<td>Two-way 2 x 2 ANOVA</td>
</tr>
<tr>
<td>Lose-shift</td>
<td>None</td>
<td>Negative study (Negative-neutral vs. negative-negative)</td>
<td>Two-way 2 x 2 ANOVA</td>
</tr>
<tr>
<td>Cross-correlation for each block</td>
<td>5 Lags</td>
<td>Experiment</td>
<td>Three-way 5 x 2 x 2 mixed-ANOVA</td>
</tr>
</tbody>
</table>

The lab-based questionnaires were compared between groups and experiments using 2 x 2 ANOVAs with the factors Group (checker vs. controls) and Experiment for each study. The exception was the strategy questionnaires that was analysed by using a Pearson Chi-square or Fisher’s exact test to compare groups for each experiment and experiments for each group. No interaction effects were investigated for the strategy questionnaire. Findings for these lab-based questionnaires will be discussed together with Chapter 5 and 7 findings in Chapter 8 (General discussion), except for the strategy questionnaires that will be also discussed here in Chapter 6.4.
Finally, Kendall-tau correlations were computed for each experiment separately, considering only checkers to investigate if (1) if symptom strength recorded with OCI-R total and OCI-R checking subscale scores, (2) comorbidities recorded with the DASS, and (3) measures from the laboratory-based questionnaires are related to behavioural measures recorded in the binary decision making task. These measures are: Accuracies for block 1, block 2, and block 3; win-shift and lose-shift; cross-correlations for lags 1-5 for block 1, 2 and 3.

6.3. Results

6.3.1. Positive study (study 2 as Positive-neutral vs. study 4 as positive-positive)

6.3.1.1. Picture Validation task

6.3.1.1.2. Valence scale

The mixed 4 x 2 x 2 ANOVA with the factors Picture Type, Group and Experiment revealed a significant main effect of Picture Type for the valence ratings, $F(2.28,96.04) = 358.1, p < .001, \eta^2 = .89$, but no other significant main effect or interaction. The related post-hoc paired t-tests were all statistically significant (all $t(45) > 3.50$, all p’s < .01). As expected from previous results in Chapters 3 and 5, positive pictures were rated as more positive than neutral, checking-related and negative pictures; neutral pictures were rated more positive than checking-related and negative pictures; and checking-related pictures were related as less negative than negative pictures.

6.3.1.1.3. Arousal scale

ANOVA results showed only a significant main effect of Picture Type for the arousal ratings, $F(1.83,77.03)=25.14, p < .001$. Pairwise comparisons were conducted to investigate this effect but not all pairwise comparisons were statistically significant and they are shown in table 6.5. Basically, results showed that positive, negative and checking-related pictures generated more arousal than neutral pictures. Positive and negative pictures and positive and checking-related provoked similar arousal ratings. However, checking-related pictures provoked less arousal than negative pictures. No other significant effect was found.
Table 6.5.

Pairwise comparisons for picture type in the picture validation task in the arousal scale for the positive study

<table>
<thead>
<tr>
<th>Picture-type (Arousal scale)</th>
<th>Mean and standard error</th>
<th>T-test</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive vs. Neutral</td>
<td>4.22±0.19 vs. 5.38±0.09</td>
<td>( t(45) = 6.30 ), ( p &lt; .001^{**} )</td>
<td>( r = .68 )</td>
</tr>
<tr>
<td>Positive vs. Negative</td>
<td>4.22±0.19 vs. 3.93±0.12</td>
<td>( t(45) = 1.21 ), ( p = .23 )</td>
<td></td>
</tr>
<tr>
<td>Positive vs. Checker-related</td>
<td>4.22±0.19 vs. 4.56±0.10</td>
<td>( t(45) = 1.46 ), ( p = .15 )</td>
<td></td>
</tr>
<tr>
<td>Neutral vs. Negative</td>
<td>5.38±0.09 vs. 3.93±0.12</td>
<td>( t(45) = 9.48 ), ( p &lt; .001^{**} )</td>
<td>( r = .81 )</td>
</tr>
<tr>
<td>Neutral vs. Checker-related</td>
<td>5.38±0.09 vs. 4.56±0.10</td>
<td>( t(45) = 7.14 ), ( p &lt; .001^{**} )</td>
<td>( r = .72 )</td>
</tr>
<tr>
<td>Negative vs. Checker-related</td>
<td>3.93±0.12 vs. 4.56±0.10</td>
<td>( t(45) = 5.66 ), ( p &lt; .001^{**} )</td>
<td>( r = .64 )</td>
</tr>
</tbody>
</table>

6.3.1.1.4. Comfort scale

Results showed there was a significant main effect of Picture Type for the comfort ratings, \( F(2.49,104.92) = 300.18, p < .001, \eta^2 = .87 \). As expected, all post-hoc paired t-tests were statistically significant (all \( t(45) > 3.50 \), all \( p’s < .01 \)), as expected. Basically, positive pictures were rated as more comfortable than neutral, checker-related and negative pictures; neutral pictures were rated as more comfortable than negative and checker-related pictures and checker-related pictures as more comfortable than negative pictures. No other significant effect was found.

In summary, the findings from the picture validation task reported in Chapters 3 and 5 were only partly confirmed in this study. In Chapter 3, group differences were reported for the valence ratings of the checking-related pictures and the comfort ratings of the negative and checking-related pictures. In Chapter 5, group differences were only found for the comfort ratings of the checking-related pictures. Here, no significant group differences were found for the picture ratings in this study. This will be discussed in Chapter 8 (General Discussions).
6.3.1.2. Binary decision making task

6.3.1.2.1. Accuracy

For each block, a two-way between-participant 2 x 2 ANOVAs with the factors Group (checker vs. non-checker) and Experiment (positive-neutral vs. positive-positive experiment) was conducted to investigate group and experiment differences for each block.

The accuracy analysis of block 1 revealed a marginally significant main effect, \( F(1,.42) = 4.02, p = .051, \eta^2 = .08, \) showing that the mean accuracy was numerically but not statistically higher for the positive-positive experiment (0.58 ± 0.01) compared to the positive-neutral experiment (0.56± 0.01). There were no significant main effect of Group and no significant interaction. The accuracy analysis of blocks 2 and 3 did not reveal any significant effects.

6.3.1.2.4. Sub-block Accuracy (Learning curves)

A step further was taken by looking at sub-block accuracy levels to explore potentially learning effects over the duration of a block, i.e. learning curves. These learning curves are displayed in Figure 6.1, considering all participants and positive experiments together for each block and it seems that accuracy improves over time for all blocks. We also investigated group and experimental differences in these learning curves.
For this, three-way 5 x 2 x 2 mixed-ANOVAs were conducted for each block. There was significant main effects of Sub-block for block 1, $F(4.,168) = 8.91$, $p < .001$, $\eta^2_p = .17$, block 2, $F(4.,168) = 6.36$, $p < .001$, $\eta^2_p = .13$, and block 3, $F(3.51.,168) = 17.71$, $p < .001$, $\eta^2_p = .29$. Related post-hoc paired t-tests are displayed in Table 6.6.

These post-hoc tests revealed that participants were able to increase performance in all the three blocks; as accuracy levels progressively increased across sub-blocks. However, learning in block 1 was delayed relative to learning in blocks 2 and 3. In block 1, participants needed at least two sub-blocks to increase accuracy (from sub-blocks 1 to 3 and again from 3 to 5), whilst they only needed the first sub-block (from sub-block 1 to 2) to learn the task in blocks 2 and 3 (see Figure 6.1), which could be related to the difficulty levels of each block but also to the fixed block order. These results confirmed findings from Chapter 5.

For all blocks, all other main effects and the interaction between Experiment and Sub-block, the interaction between Group and Sub-block, and the interaction between Experiment, Group and Sub-block were not significant.

In summary, the accuracy analysis for each block revealed: Firstly, learning occurred in all three blocks but it was slower in block 1. Secondly, there were no group or
experimental differences for the random sequence block (block 1), for the pattern sequence block (block 2), and for block 3 where target option maximisation and pattern search strategies were combined. These findings will be discussed later on.
Table 6.6
Pairwise comparisons for the main effect of sub-block accuracy for blocks 1, 2 and 3, considering all participants and experiments together

<table>
<thead>
<tr>
<th>Sub-block accuracy</th>
<th>Mean and standard error</th>
<th>T-test</th>
<th>Effect size</th>
<th>Mean and standard error</th>
<th>T-test</th>
<th>Effect size</th>
<th>Mean and standard error</th>
<th>T-test</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Block 1</td>
<td></td>
<td></td>
<td>Block 2</td>
<td></td>
<td></td>
<td>Block 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub-block 1 vs. Sub-block 2</td>
<td>0.52±0.02 vs. 0.54±0.01</td>
<td>t(45) = 0.98, p = .32</td>
<td>r = .45</td>
<td>0.57±0.02 vs. 0.63±0.02</td>
<td>t(45) = 0.02, p = .02*</td>
<td>r = .42</td>
<td>0.59±0.02 vs. 0.70±0.02</td>
<td>t(45) = .94, p = .32</td>
<td>r = .32</td>
</tr>
<tr>
<td>Sub-block 1 vs. Sub-block 3</td>
<td>0.52±0.02 vs. 0.58±0.01</td>
<td>t(45) = 2.53, p = .46</td>
<td>r = .46</td>
<td>0.57±0.02 vs. 0.64±0.02</td>
<td>t(45) = 0.02, p = .02*</td>
<td>r = .42</td>
<td>0.59±0.02 vs. 0.70±0.02</td>
<td>t(45) = 1.70, p = .42</td>
<td>r = .42</td>
</tr>
<tr>
<td>Sub-block 2 vs. Sub-block 1</td>
<td>0.54±0.01 vs. 0.57±0.01</td>
<td>t(45) = 1.54, p = .32</td>
<td>r = .32</td>
<td>0.63±0.02 vs. 0.66±0.02</td>
<td>t(45) = 0.02, p = .02*</td>
<td>r = .32</td>
<td>0.70±0.02 vs. 0.70±0.02</td>
<td>t(45) = 1.44, p = .25</td>
<td>r = .25</td>
</tr>
<tr>
<td>Sub-block 2 vs. Sub-block 2</td>
<td>0.54±0.01 vs. 0.57±0.01</td>
<td>t(45) = 1.54, p = .32</td>
<td>r = .32</td>
<td>0.63±0.02 vs. 0.66±0.02</td>
<td>t(45) = 0.02, p = .02*</td>
<td>r = .32</td>
<td>0.70±0.02 vs. 0.70±0.02</td>
<td>t(45) = 1.44, p = .25</td>
<td>r = .25</td>
</tr>
<tr>
<td>Sub-block 3 vs. Sub-block 1</td>
<td>0.58±0.01 vs. 0.64±0.02</td>
<td>t(45) = 2.53, p = .46</td>
<td>r = .46</td>
<td>0.64±0.02 vs. 0.67±0.02</td>
<td>t(45) = 0.02, p = .02*</td>
<td>r = .42</td>
<td>0.70±0.02 vs. 0.70±0.02</td>
<td>t(45) = 1.44, p = .25</td>
<td>r = .25</td>
</tr>
<tr>
<td>Sub-block 3 vs. Sub-block 2</td>
<td>0.58±0.01 vs. 0.64±0.02</td>
<td>t(45) = 2.53, p = .46</td>
<td>r = .46</td>
<td>0.64±0.02 vs. 0.67±0.02</td>
<td>t(45) = 0.02, p = .02*</td>
<td>r = .42</td>
<td>0.70±0.02 vs. 0.70±0.02</td>
<td>t(45) = 1.44, p = .25</td>
<td>r = .25</td>
</tr>
<tr>
<td>Sub-block 4 vs. Sub-block 1</td>
<td>0.57±0.01 vs. 0.62±0.01</td>
<td>t(45) = 1.54, p = .32</td>
<td>r = .32</td>
<td>0.63±0.02 vs. 0.66±0.02</td>
<td>t(45) = 0.02, p = .02*</td>
<td>r = .32</td>
<td>0.70±0.02 vs. 0.70±0.02</td>
<td>t(45) = 1.44, p = .25</td>
<td>r = .25</td>
</tr>
<tr>
<td>Sub-block 4 vs. Sub-block 2</td>
<td>0.57±0.01 vs. 0.62±0.01</td>
<td>t(45) = 1.54, p = .32</td>
<td>r = .32</td>
<td>0.63±0.02 vs. 0.66±0.02</td>
<td>t(45) = 0.02, p = .02*</td>
<td>r = .32</td>
<td>0.70±0.02 vs. 0.70±0.02</td>
<td>t(45) = 1.44, p = .25</td>
<td>r = .25</td>
</tr>
<tr>
<td>Sub-block 4 vs. Sub-block 3</td>
<td>0.57±0.01 vs. 0.62±0.01</td>
<td>t(45) = 1.54, p = .32</td>
<td>r = .32</td>
<td>0.63±0.02 vs. 0.66±0.02</td>
<td>t(45) = 0.02, p = .02*</td>
<td>r = .32</td>
<td>0.70±0.02 vs. 0.70±0.02</td>
<td>t(45) = 1.44, p = .25</td>
<td>r = .25</td>
</tr>
<tr>
<td>Sub-block 5 vs. Sub-block 1</td>
<td>0.62±0.01 vs. 0.64±0.02</td>
<td>t(45) = 1.54, p = .32</td>
<td>r = .32</td>
<td>0.63±0.02 vs. 0.66±0.02</td>
<td>t(45) = 0.02, p = .02*</td>
<td>r = .32</td>
<td>0.70±0.02 vs. 0.70±0.02</td>
<td>t(45) = 1.44, p = .25</td>
<td>r = .25</td>
</tr>
<tr>
<td>Sub-block 5 vs. Sub-block 2</td>
<td>0.62±0.01 vs. 0.64±0.02</td>
<td>t(45) = 1.54, p = .32</td>
<td>r = .32</td>
<td>0.63±0.02 vs. 0.66±0.02</td>
<td>t(45) = 0.02, p = .02*</td>
<td>r = .32</td>
<td>0.70±0.02 vs. 0.70±0.02</td>
<td>t(45) = 1.44, p = .25</td>
<td>r = .25</td>
</tr>
<tr>
<td>Sub-block 5 vs. Sub-block 3</td>
<td>0.62±0.01 vs. 0.64±0.02</td>
<td>t(45) = 1.54, p = .32</td>
<td>r = .32</td>
<td>0.63±0.02 vs. 0.66±0.02</td>
<td>t(45) = 0.02, p = .02*</td>
<td>r = .32</td>
<td>0.70±0.02 vs. 0.70±0.02</td>
<td>t(45) = 1.44, p = .25</td>
<td>r = .25</td>
</tr>
<tr>
<td>Sub-block 5 vs. Sub-block 4</td>
<td>0.62±0.01 vs. 0.64±0.02</td>
<td>t(45) = 1.54, p = .32</td>
<td>r = .32</td>
<td>0.63±0.02 vs. 0.66±0.02</td>
<td>t(45) = 0.02, p = .02*</td>
<td>r = .32</td>
<td>0.70±0.02 vs. 0.70±0.02</td>
<td>t(45) = 1.44, p = .25</td>
<td>r = .25</td>
</tr>
</tbody>
</table>
6.3.1.2.5. **Exploration / exploitation strategy: Win-stay / lose-shift in Block 1**

Figure 6.2 indicates that checkers performed more exploitation in the positive-positive experiment than in the positive-neutral experiment, while non-checkers seem to use similar strategies in both experiment. In fact, this figure indicates that checkers and non-checkers are not doing excessive exploration (more win-shift and more lose-shift) because only one participant presented a higher probability of shifting after losing and winning in the positive-positive experiment (see right superior quadrant for both experiments in figure 6.2). Besides, it is possible to see that both groups prefer to shift after losing, as only one participant is seen at the right inferior quadrant of figure 6.2 (more win-shift/lose-stay) in the positive-neutral experiment, where the probability of shifting after losing is the lowest. Additionally, checkers seem to be doing more shifts after losing than non-checkers in the positive-neutral experiment, while checkers seem to be doing less shifts after losing in the positive-positive experiment in comparison to non-checkers (see left superior and inferior quadrant of figure 6.2).

![Figure 6.2 Win-shift / Lose-shift probability per participant divided by group and experiment for the positive study. The left and inferior quadrant of the figure indicates more exploitation after winning and after losing (win-stay / lose-stay). The right and superior quadrant indicates more exploration after winning and losing (win-shift / lose-shift). The left and superior](image-url)

**Figure 6.2 Win-shift / Lose-shift probability per participant divided by group and experiment for the positive study. The left and inferior quadrant of the figure indicates more exploitation after winning and after losing (win-stay / lose-stay). The right and superior quadrant indicates more exploration after winning and losing (win-shift / lose-shift). The left and superior**
quadrant indicates *more exploration after losing only* (win-stay / lose-shift). The right and inferior quadrant indicates *more exploration after winning only* (win-shift / lose-stay). The right figure indicates that checkers are doing less exploration in the positive-positive experiment, as they present a lower probability of shifting after winning and losing.

Results from the two-way 2 x 2 ANOVA confirmed these differences and revealed that there was a marginal main effect of experiment, (Mean (Positive-neutral): 0.29±0.03; Mean (Positive-positive): 0.22±0.02; $F(1, 42) = 3.36, p = .07, \eta^2 = .07$) for win-shift, where participants shifted more after winning in the previous trial in the positive-neutral experiment than in the positive-positive experiment. There was no significant main effect of group and no significant interaction between experiment and group for win-shift.

For lose-shift, there were no significant main effects of Group and Experiment but a significant interaction between the factors Group and Experiment was present, $F(1, 42) = 15.37, p < .001, \eta^2 = .26$. This interaction was explored further with four independent t-tests where groups were compared for each experiment separately and where experiments were compared for each group (figure 6.3). There was a *marginally* significant group difference for the *positive-neutral* experiment (Mean (Checkers): 0.56 ± 0.03; Mean (Non-checkers): 0.45± 0.04; $t(20)= 2.25, p_b = .07, r=0.44$), where checkers shifted more often after losing than non-checkers (Figure 6.3). Additionally, there was a significant effect of group for the *positive-positive* experiment (Mean (Checkers): 0.35 ± 0.05; Mean (Non-checkers): 0.53± 0.03; $t(22)= 3.28, p_b < .001, r = .57$), where checkers shifted less often after losing than non-checkers (Figure 6.3).

Furthermore, *checkers* shifted more after losing in the positive-neutral experiment than in the positive-positive experiment (Mean (Positive-neutral): 0.56 ± 0.03; Mean (Positive-positive): 0.35 ± 0.05; $t(20)= 3.82, p_b < .001, r = .64$), but there was no significant difference between experiments for the *non-checker group* (Figure 6.3).
Figure 6.3 Lose-shift mean values for checkers and non-checkers in block 1 between the positive-neutral and the positive-positive experiment in the positive study. Figure indicates that checkers shifted more after losing in the positive-neutral than in the positive-positive experiment and that checkers shifted less after losing than non-checkers in the positive-positive experiment.

In summary, all participants shifted less after winning in the positive-positive experiment compared to the positive-neutral experiment but this is possibly related with the fact that checkers shifted less in this experiment, meaning all participants were influenced by the reward magnitude manipulation by increasing the exploitation of the environment after winning in the positive-positive experiment. The findings for the lose-shift measure are more complex. Here, checkers shifted less after losing in the positive-positive experiment compared to the positive-neutral experiment whereas lose-shift values did not change for non-checkers. Hence, lose-shift strategy changes towards exploitation were only found in checkers but not in non-checkers.
6.3.1.2.6. Cross-correlation

6.3.1.2.6.1. Cross-correlation for Block 1

Figure 6.4 shows the auto- and cross-correlations for block 1 in the positive study, considering groups and experiments together. Here, the difference between auto-and cross-correlation scores seems to be most evident for lag 1, but maybe also present for lag 3. The three-way 5 x 2 x 2 mixed ANOVA on the auto- and cross-correlation difference scores revealed a significant main effect of Lag, $F(4.,168) = 15.85$, $p < .001$, $\eta^2 = .27$, as in Chapter 5. This was caused by a significantly higher difference between the participant’s auto- and cross-correlation values for lag 1 than for lags 2, 3, 4, and 5 (all $t(45) > 3.50$, all $p$’s < .01). This might be because participants assumed a greater dependency between the current and the previous event (lag 1) than present in a random sequence. More specifically, participants used the previous trial to make decisions and automatically chose the best option more often than the worst option in the task, thereby, increasing the cross-correlation value for lag 1 and also augmenting the difference between the participants and experimental sequence correlation value. In addition, the difference between auto-and cross-correlations was higher at lag 3 compared to lag 2 (Mean (lag 2): -0.02 ± 0.03; Mean (lag 3): 0.07± 0.02; $t(45) = 3.93$, $p_b < .001$, $r=0.50$), and was higher at lag 3 compared to lag 5 (Mean (lag 3): -0.07 ± 0.02; Mean (lag 5): 0.004± 0.02; $t(45) = 3.45$, $p_b = .01$, $r=0.45$). Again, this effect is more difficult to explain and it might be more due to uncontrolled irregularities in the experimental sequences, because the experimental sequence score is slightly enhanced at lag 3 and because this difference was not found for the others pairwise comparisons of lag 2 (2-4, 2-5) or lag 3 (3-4, 3-5). In this way, supporting findings from Chapter 5, participants usually employ the immediate previous results to make predictions in a random sequence.
Figure 6.4 Auto- and cross-correlation values for the experimental sequence and the participant’s sequence for lags 1 to 5 in block 1 for the positive study. The graphic combines data from both groups and both experiments. Figure indicates that participants cross-correlated more in lag 1 than it was expected when considering the sequence’s values of correlation in lag 1.

The three-way ANOVA 5 x 2 x 2 for block 1 also showed a significant interaction of the factors Lag, Experiment and Group, \( F(4,168) = 3.38, p = .01, \eta^2 = .07 \). For this interaction, five distinct two-way 2 x 2 ANOVAs were conducted for each lag in block 1. Results showed there was a significant Group x Experiment interaction for lag 1, \( F(1,42) = 4.15, p = .05, \eta^2 = .09 \). Independent t-tests were conducted to explore this interaction showing a significantly lower cross-correlation for checkers compared to non-checkers for the lag 1 in the positive-positive experiment, represented by the hashed lines (Mean (Checkers): -0.04 ± 0.05; Mean (Non-checkers): 0.19 ± 0.03; \( t(22) = 2.25, p_b = .03, r = .48 \)), whereas there was no group difference for the positive-neutral experiment (Figure 6.5). Note, there were no significant effects for lags 2, 3, 4 and 5.
Figure 6.5 Auto- and cross-correlation values for the experimental sequence and the participant’s sequence for lags 1 to 5 in block 1 in the positive study. Comparison between checkers and non-checkers for the positive-neutral (solid lines) and the positive-positive experiment (hashed lines). Figure shows that checkers presented a lower cross-correlation value than non-checkers for lag 1 in the positive-positive experiment.

6.3.1.2.6.1.2. Cross-correlation for Block 2

Figure 6.6 shows that both groups, checkers and non-checkers, were able to follow the pattern of the experimental event sequence in this block, when considering both positive experiments together. As expected, the absolute cross-correlation values were not as large as the absolute auto-correlation values because participants could not follow the pattern perfectly. This can be, at least partly, explained by the added probabilistic noise to the sequence. In addition, Figure 6.6 did not reveal group differences in the participant responses to the pattern sequence.
Figure 6.6 Cross-correlation probability from lags 1 to 5 between checkers and non-checkers considering both experiments in block 2 for the positive study. The averaged correlation values of the experimental sequence are shown in black. Figure shows that checkers and non-checkers presented similar cross-correlation values, even though there was a marginal significance between groups for lag 3.

This was reflected in the findings from the statistical analysis. The three-way 5 x 2 x 2 ANOVA for block 2 revealed a significant main effect of Lag, \( F(1.88, 79.02) = 72.67, p < .001, \eta^2_p = .63 \), which was expected because of the presence of an experimental pattern sequence. In addition, a marginally significant interaction for the factors Lag and Group was found, considering all experiments together, \( F(1.88, 79.02) = 3.07, p = .055, \eta^2_p = .06 \). This was explored with five independent post-hoc t-tests comparing both groups for lags 1 to 5, considering both experiments together. Results showed groups were marginally significantly different for lag 3 (Mean (Checkers): -0.08 ± 0.03; Mean (Non-checkers): 0.04± 0.03; \( t(44) = 2.77, p_b = .05, r=0.38 \)), where checkers were more distant from the auto-correlation value than non-checkers, indicating that non-checkers learned the pattern structure of the experimental sequence better than checkers (Figure 6.6). Note, there were no other significant effects. These data could fit to the findings from Chapter 5 where non-checkers were closer to the experimental sequence in the positive-neutral experiment (0.05±0.02) than in the negative-neutral (0.18±0.03), \( t(21) = 2.77, p_b = .02 \). The later numerical trend was similar,
when considering both positive experiments together, even though the difference was only marginally significant.

### 6.3.1.2.6.1.3. Cross-correlation for Block 3

Figure 6.7 shows cross-correlation values for block 3 in the positive study, when considering differences between groups and experiments. The three-way ANOVA $5 \times 2 \times 2$ for block 3 showed a main effect of Lag, $F(1.72,72.50) = 57.52$, $p < .001$, $\eta^2 = .57$, which was expected as lag values were different in the pattern sequence. No other effects were significant (figure 6.7).

![Figure 6.7 Cross-correlation probability from lags 1 to 5 between checkers and non-checkers for the positive-neutral and positive-positive experiments in block 3. The averaged correlation values of the experimental sequence are shown in black. Figure shows there was no significant difference between groups and experiments.](image)

In summary, the findings from the auto- and cross-correlation analysis showed that, as expected, participants had enhanced cross-correlation values in lag 1 of the random-sequence block (block 1) and that they could follow the pattern sequences in blocks 2 and 3. There were two group differences in the cross-correlation data. Firstly, in accordance with the results from the lose-shift analysis of block 1, non-checkers had higher cross-correlations in lag 1 than checkers in the positive-positive experiment whereas no group differences were found for the positive-neutral experiment. Secondly, marginal group differences were also
revealed for block 2 (conditional probability manipulation). Here, non-checkers were marginally closer to the experimental sequence than checkers for lag 3 when considering both experiments together. Finally, there was no significant effect of group or experimental differences for block 3 (target option and conditional probability manipulation). These findings will be discussed in Chapter 6.4

6.3.1.3. Laboratory-based questionnaires

6.3.1.3.1. Intolerance of uncertainty scale (IUS)

The two-way 2 x 2 ANOVA analysis with the factors Group and Experiment revealed that checkers had higher intolerance of uncertainty scores compared to non-checkers. More specifically, there were significant main effects of Group for factor 1 (Uncertainty has negative behavioural and self-referent implications; Mean (checkers): 35.83 ± 1.63; Mean (non-checkers): 30.65 ± 1.38; $F(1,42) = 5.59, p = .02, \eta^2 = .11$), factor 2 (uncertainty is unfair and spoils everything; Mean (checkers): 35.52 ± 1.68; Mean (non-checkers): 30.65 ± 1.07; $F(1,42) = 5.57, p = .02, \eta^2 = .11$) and the total score (Mean (checkers): 71.35 ± 2.90; Mean (non-checkers): 61.03 ± 2.18; $F(1,42) = 7.25, p = .01, \eta^2 = .14$). This supports findings from Chapter 5, where checkers also were more intolerant to uncertainty than non-checkers.

6.3.1.3.2. Domain Specific Risk-Taking Scale and Risk-Perception scale (DOSPERT)

The two-way 2 x 2 ANOVA analysis found no significant main effects and interactions for the financial risk taking and risk perception subscale scores and the total risk taking and risk perception score. Thus, it seems checkers and non-checkers presented similar values of risk taking in both experiments.

6.3.1.3.3. UPPS Impulsivity Scale

The two-way 2 x 2 ANOVA analysis found no significant main effects and interactions for all the UPPS sub-sections and total score. Thus, it seems checkers and non-checkers had similar impulsivity ratings in both experiments.

6.3.1.3.4. South Oaks Gambling Screen

The two-way 2 x 2 ANOVA analysis showed a significant main effect of Group for gambling, (Mean (checkers): 1.57 ± 0.38; Mean (non-checkers): 0.3 ± 0.22; $F(1,42) = 8.09$,
p = .007, \eta^2 = .16), meaning that checkers showed more gambling behaviour than non-checkers.

6.3.1.3.5. **Strategy questionnaires**

Categorical data from the strategy questionnaire showed that checkers and non-checkers seem to employ similar strategies and visualize patterns in the positive-neutral and in the positive-positive experiment in blocks 1 and 3.

However, checkers saw more patterns in block 2 in the positive-positive experiment than in the positive-neutral experiment (Positive-neutral): 4 Yes; 7 No (Positive-positive): 10 Yes; 2 No, p = .03). There was also a similar but marginally significant effect of experiment for non-checkers ((Positive-neutral): 6 Yes; 5 No, (Positive-positive): 11 Yes; 1 No, p = .07), i.e. non-checkers saw numerically more patterns in the positive-positive than in the positive-neutral experiment.

In summary, the analysis of the lab-based questionnaires revealed that checkers showed greater levels of intolerance of uncertainty and they were bigger gamblers than non-checkers. In addition, checkers used more strategies in block 2 in the positive-positive experiment than in the positive neutral experiment. Non-checkers presented the same tendency but this effect was marginally significant. There were no strategy differences between groups and experiments for blocks 1 and 3.

6.3.1.4. **Correlation analysis between behavioural measures from the binary decision making task and laboratory-based questionnaire scores**

6.3.1.4.1. **Correlation analysis for the positive-neutral experiment**

This analysis was already reported in Chapter 5. As a reminder, the Kendall Tau correlation analysis revealed the following findings. Firstly, symptom severity in checkers affected performance in the binary decision making task, i.e. symptom severity (OCI-R total) was positively correlated with the accuracy level in block 1, \( r_\tau = .50, p = .03 \), and block 3, \( r_\tau = .54, p = .02 \). Secondly, comorbidity levels correlated with the binary decision making task performance. More specifically, higher depression levels in checkers were related to enhanced accuracy levels in the binary decision making task in block 1, \( r_\tau = .61, p = .01 \) and higher anxiety levels in checkers were related to higher lose-shift scores in block 1, \( r_\tau = .51, p \)
Finally, laboratory-based questionnaire ratings did not correlate with the task performance measures in the binary decision making task.

6.3.1.4.2. Correlation analysis for the positive-positive experiment

For the positive-positive experiment, enhanced symptom severity was not significantly correlated with any binary decision making variables. Secondly, higher comorbidity levels correlated with the binary decision making task performance. More specifically, higher scores of anxiety in checkers were related to lower accuracy in block 3, \( r_t = .50, p = .02 \). Thirdly, laboratory-based questionnaire ratings did not correlate with task performance. Finally, individual positive valence scores did not correlate with the binary decision making task performance in the positive-positive experiment.

In summary, there were no common correlations between the positive-neutral and the positive-positive experiment which makes the interpretation of the correlational findings rather difficult. Correlation interpretations will be treated with utmost caution because of this. For example, it seems that while symptom strength influences accuracy of block 1 in the positive-neutral experiment, symptom strength does not influence accuracy in the positive-positive experiment in block 1. Another interesting finding from the correlational analysis is that there were no significant correlations between intolerance of uncertainty or gambling behaviour ratings, which are subjective self-rating measures, and the binary decision making task performance in this study. This will be discussed later on in Chapter 8.

6.3.2. Negative study (Negative-neutral as study 3 vs. negative-negative as study 5)

6.3.2.1. Picture Validation task

6.3.2.1.1. Valence scale

The mixed 4 x 2 x 2 ANOVA with the factors Picture Type, Group and Experiment revealed a significant main effect of Picture Type for the valence ratings \( F(2.67,117.67) = 417.93, p < .001, \eta^2 = .90 \). The related post-hoc paired t-tests were all statistically significant (all \( t(47) > 3.50, \) all \( p’s < .01 \)). As expected from previous results in Chapters 3 and 5 and in Chapter 6 (Positive Study), positive pictures were rated as more positive than neutral, checking-related and negative pictures; neutral pictures were rated as more positive than
checking-related; and negative pictures and checking-related pictures were rated as less negative than negative pictures. No other effects were significant.

6.3.2.1.2. Arousal scale

The mixed ANOVA results only showed a significant main effect of Picture Type on arousal ratings, $F(1.70,75.03)=16.08$, $p < .001$. To investigate the significant main effect of picture type, 6 pairwise comparisons were conducted. Not all pairwise comparisons were statistically significant and they are shown in table 6.7. Basically, they showed that positive, negative and checking-related pictures generated more arousal than neutral pictures. Positive and negative pictures and positive and checking-related presented similar amounts of arousal. Checking-related pictures, however, presented less arousal than negative pictures, considering both experiments and groups.

Table 6.7.

Pairwise comparisons for picture type in the picture validation task in the arousal scale for the negative study

<table>
<thead>
<tr>
<th>Picture-type (Arousal scale)</th>
<th>Mean and standard error</th>
<th>T-test</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive vs. Neutral</td>
<td>4.35±0.23 vs. 4.45±0.13</td>
<td>$t(47) = 5.28$, $p &lt; .001**$</td>
<td>$r = .61$</td>
</tr>
<tr>
<td>Positive vs. Negative</td>
<td>4.35±0.23 vs. 4.06±0.20</td>
<td>$t(47) = 0.95$, $p = .34$</td>
<td></td>
</tr>
<tr>
<td>Positive vs. Checker-related</td>
<td>4.35±0.23 vs. 4.61±0.16</td>
<td>$t(47) = 1.02$, $p_b = .31$</td>
<td></td>
</tr>
<tr>
<td>Neutral vs. Negative</td>
<td>4.45±0.13 vs. 4.06±0.20</td>
<td>$t(47) = 7.22$, $p &lt; .001**$</td>
<td>$r = .72$</td>
</tr>
<tr>
<td>Neutral vs. Checker-related</td>
<td>4.45±0.13 vs. 4.61±0.16</td>
<td>$t(47) = 7.37$, $p &lt; .001**$</td>
<td>$r = .73$</td>
</tr>
<tr>
<td>Negative vs. Checker-related</td>
<td>4.06±0.20 vs. 4.61±0.16</td>
<td>$t(47) = 3.80$, $p_b &lt; .001**$</td>
<td>$r = .48$</td>
</tr>
</tbody>
</table>
6.3.2.1.3. Comfort scale

Results showed there was a significant main effect of Picture Type on comfort ratings, $F(2.91,128.21) = 270.33, p < .001, \eta^2 = .86$. As expected, all post-hoc paired t-tests were statistically significant (all $t(47) > 3.50$, all $p$’s < .01). Basically, positive pictures were rated as more comfortable than neutral, checking-related and negative pictures; neutral pictures were rated as more comfortable than negative and checking-related pictures; and checking-related pictures were rated as more comfortable than negative pictures. No other significant effect was found.

In summary, the findings from the picture validation task study reported in Chapters 3 and 5 were only partly confirmed in this study. As in Chapter 6 (Positive study), there was no group differences for the comfort scale and valence scale, regarding checking-related pictures or negative pictures. These rating differences between studies will be discussed in Chapter 8 (General Discussions).

6.3.2.2. Binary decision making task

6.3.2.2.1. Accuracy

For each block, a two-way between-participant 2 x 2 ANOVA with the factors Group (checker vs. non-checker) and Experiment (negative-neutral vs. negative-negative experiment) was conducted to investigate group and experiment differences.

The analyses of block 1 and block 3 did not reveal any significant main effects or interaction effects.

In contrast, the analysis of block 2 revealed a significant interaction between the factors Group and Experiment, $F(1.,44) = 4.71, p = .03, \eta^2 = .9$. Post-hoc independent t-tests showed that checkers (0.68 ± 0.03; $t(22) = 2.61, p_b=0.03, r=-.48$) had significantly higher accuracy levels than non-checkers (0.55± 0.03; $t(22) = 2.61, p_b=0.03, r=-.48$) in the negative-neutral experiment, whereas no significant group differences were found for the negative-negative experiment (Mean (Checker): 0.62 ± 0.03; Mean (Non-checker): 0.63± 0.02; $t(22) = 0.21, p_b = .83$). There was no significant difference between experiments for checkers and for non-checkers.

6.3.2.2.2. Sub-block Accuracy (Learning curves)

A step further was taken by looking at sub-block accuracy levels to explore potentially learning effects over the duration of a block. i.e. learning curves. These learning
curves are displayed in Figure 6.8, showing values for both groups and experiments together. It seems that accuracy improves over time for all blocks. We also investigated group and experimental differences in these learning curves.
Figure 6.8 Sub-block accuracy mean values for all participants and experiments together for blocks 1, 2 and 3 for the Negative study. Figure indicates that learning improved across sub-blocks.

Indeed, a three-way 5 x 2 x 2 mixed-ANOVA was conducted for each block and confirmed these predictions. They showed significant main effects of Sub-block for block 1, $F(4.,176) = 14.47$, $p < .001$, $\eta^2 = .24$; block 2, $F(4.,176) = 2.66$, $p < .001$, $\eta^2 = .05$ and block 3, $F(3.58.,157.72) = 17.71$, $p < .001$, $\eta^2 = .15$. Post-hoc paired t-tests for each sub-block are displayed in Table 6.8. There was no significant interaction between experiment and sub-block accuracy for blocks 1, 2 and 3, no significant interaction between group and sub-block accuracy for blocks 1, 2 and 3 and no significant interaction between experiment, group and sub-block for all the blocks.
Table 6.8

Pairwise comparisons for the main effect of sub-block accuracy for blocks 1, 2 and 3, considering groups and experiments together

<table>
<thead>
<tr>
<th>Sub-block accuracy</th>
<th>Mean and standard error</th>
<th>T-test</th>
<th>Effect size</th>
<th>Mean and standard error</th>
<th>T-test</th>
<th>Effect size</th>
<th>Mean and standard error</th>
<th>T-test</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Block 1</td>
<td>Block 2</td>
<td>Block 3</td>
<td>Block 1</td>
<td>Block 2</td>
<td>Block 3</td>
<td>Block 1</td>
<td>Block 2</td>
<td>Block 3</td>
</tr>
<tr>
<td>Sub-block 1 vs. Sub-block 2</td>
<td>0.50±0.01 vs. 0.55±0.01</td>
<td>( t(47) = 2.94, p_b = .05^* )</td>
<td>( r = .39 )</td>
<td>0.55±0.02 vs. 0.64±0.02</td>
<td>( t(47) = 4.18, p_b = .05^* )</td>
<td>( r = .39 )</td>
<td>0.59±0.02 vs. 0.65±0.02</td>
<td>( t(47) = 3.22, p_b = .02^* )</td>
<td>( r = .42 )</td>
</tr>
<tr>
<td>Sub-block 1 vs. Sub-block 3</td>
<td>0.50±0.01 vs. 0.57±0.01</td>
<td>( t(47) = 3.84, p_b &lt; .001^* )</td>
<td>( r = .49 )</td>
<td>0.55±0.02 vs. 0.62±0.02</td>
<td>( t(47) = 5.58, p_b = .02^* )</td>
<td>( r = .52 )</td>
<td>0.59±0.02 vs. 0.68±0.02</td>
<td>( t(47) = 4.24, p_b &lt; .001^* )</td>
<td>( r = .40 )</td>
</tr>
<tr>
<td>Sub-block 1 vs. Sub-block 4</td>
<td>0.50±0.01 vs. 0.57±0.01</td>
<td>( t(47) = 7.55, p_b &lt; .001^* )</td>
<td>( r = .49 )</td>
<td>0.55±0.02 vs. 0.67±0.02</td>
<td>( t(47) = 3.96, p_b = .05^* )</td>
<td>( r = .52 )</td>
<td>0.59±0.02 vs. 0.65±0.02</td>
<td>( t(47) = 4.18, p_b &lt; .001^* )</td>
<td>( r = .52 )</td>
</tr>
<tr>
<td>Sub-block 1 vs. Sub-block 5</td>
<td>0.50±0.01 vs. 0.64±0.02</td>
<td>( t(47) = 9.00, p_b = .37 )</td>
<td>( r = .74 )</td>
<td>0.55±0.02 vs. 0.67±0.02</td>
<td>( t(47) = 1.27, p = .21 )</td>
<td>( r = .50 )</td>
<td>0.59±0.02 vs. 0.68±0.02</td>
<td>( t(47) = 2.22, p_b = .32 )</td>
<td>( r = .52 )</td>
</tr>
<tr>
<td>Sub-block 2 vs. Sub-block 3</td>
<td>0.57±0.01 vs. 0.55±0.01</td>
<td>( t(47) = 2.49, p_b = .16 )</td>
<td>( r = .37 )</td>
<td>0.64±0.02 vs. 0.67±0.02</td>
<td>( t(47) = 1.38, p = .17 )</td>
<td>( r = .40 )</td>
<td>0.65±0.02 vs. 0.66±0.02</td>
<td>( t(47) = 0.45, p = .26 )</td>
<td>( r = .52 )</td>
</tr>
<tr>
<td>Sub-block 2 vs. Sub-block 4</td>
<td>0.55±0.01 vs. 0.64±0.02</td>
<td>( t(47) = 4.40, p_b = .54 )</td>
<td>( r = .74 )</td>
<td>0.64±0.02 vs. 0.67±0.02</td>
<td>( t(47) = 0.38, p = .21 )</td>
<td>( r = .40 )</td>
<td>0.65±0.02 vs. 0.68±0.02</td>
<td>( t(47) = 1.88, p_b = .67 )</td>
<td>( r = .52 )</td>
</tr>
<tr>
<td>Sub-block 2 vs. Sub-block 5</td>
<td>0.55±0.01 vs. 0.64±0.02</td>
<td>( t(47) = 1.72, p_b = .02^* )</td>
<td>( r = .37 )</td>
<td>0.62±0.02 vs. 0.67±0.02</td>
<td>( t(47) = 2.46, p_b = .02^* )</td>
<td>( r = .39 )</td>
<td>0.68±0.02 vs. 0.68±0.02</td>
<td>( t(47) = 1.23, p = .22 )</td>
<td>( r = .39 )</td>
</tr>
<tr>
<td>Sub-block 3 vs. Sub-block 4</td>
<td>0.57±0.01 vs. 0.61±0.02</td>
<td>( t(47) = 3.84, p_b &lt; .001 )</td>
<td>( r = .24 )</td>
<td>0.55±0.02 vs. 0.62±0.02</td>
<td>( t(47) = 1.79, p = .24 )</td>
<td>( r = .24 )</td>
<td>0.62±0.02 vs. 0.68±0.02</td>
<td>( t(47) = 0.5, p = .68 )</td>
<td>( r = .22 )</td>
</tr>
<tr>
<td>Sub-block 3 vs. Sub-block 5</td>
<td>0.56±0.02 vs. 0.61±0.02</td>
<td>( t(47) = 1.67, p = .01 )</td>
<td>( r = .24 )</td>
<td>0.57±0.02 vs. 0.67±0.02</td>
<td>( t(47) = 1.17, p = .01 )</td>
<td>( r = .24 )</td>
<td>0.67±0.02 vs. 0.68±0.02</td>
<td>( t(47) = 1.13, p = .06 )</td>
<td>( r = .24 )</td>
</tr>
<tr>
<td>Sub-block 4 vs. Sub-block 5</td>
<td>0.64±0.02 vs. 0.61±0.02</td>
<td>( t(47) = .10 )</td>
<td>( r = .10 )</td>
<td>0.65±0.02 vs. 0.65±0.02</td>
<td>( t(47) = .25 )</td>
<td>( r = .25 )</td>
<td>0.68±0.02 vs. 0.68±0.02</td>
<td>( t(47) = .26 )</td>
<td>( r = .26 )</td>
</tr>
</tbody>
</table>
These post-hoc tests showed that participants were able learn in all three blocks, i.e. accuracy levels progressively increased across sub-blocks. Similar to previous findings, learning in block 1 was different from blocks 2 and 3. However, in this study, participants only needed one sub-block to increase accuracy in block 1 (from sub-blocks 1 to 2), as it was shown for blocks 2 and 3 for the Positive study above and Chapter 5 (study 2 vs. study 3). Nevertheless, participants continued to increase accuracy levels until the end of the block (from sub-blocks 2-5) (see Figure 6.8). In this way, this indicates that, again, block 1 might present a higher level of difficulty in comparison to the other blocks. This could be also related with the fixed block order. No other significant main effects and interactions were found.

In summary, the accuracy analysis for each block revealed: firstly, learning occurred in all three blocks but it was slower in block 1. Secondly, there were no group or experimental differences for the random sequence block 1 and for block 3 where target option maximisation and pattern search strategies were combined. However, checkers had higher accuracy levels in the pattern sequence block (block 2) in the negative-neutral experiment which was reported in Chapter 5. Interestingly, no group accuracy differences were found for block 2 for the negative-negative experiment. These findings will be discussed later.

6.3.2.2.3. Exploration / exploitation strategy: Win-stay / Lose-shift in Block 1

Figure 6.9 indicates that checkers seem to use similar strategies in the negative-negative and negative-neutral experiments, while non-checkers seem to use more exploration in the negative-negative experiment. Actually, this figure again indicates that checkers and non-checkers are not doing excessive exploration (more win-shift and more lose-shift) because only two participants presented a higher probability of shifting after losing and winning in the negative-neutral experiment (see right superior quadrant for both experiments in figure 6.9). Besides, it is possible to see that both groups prefer to shift after losing, as only one participant is seen at the right inferior quadrant of figure 6.9 (more win-shift/lose-stay) in the negative-neutral and negative-negative experiment, where the probability of shifting after losing is the lowest. Additionally, checkers seem to be doing less shifts after losing than non-checkers in the negative-negative experiment, (see left inferior quadrant of figure 6.9).
Figure 6.9 Win-shift / Lose-shift strategies per participant divided by group and negative experiments. The left and inferior quadrant of the figure indicates more exploitation after winning and after losing (win-stay / lose-stay). The right and superior quadrant indicate more exploration after winning and losing (win-shift / lose-shift). The left and superior quadrant indicates more exploration after losing only (win-stay / lose-shift). The right and inferior quadrant indicates more exploration after winning only (win-shift / lose-stay). The right figure indicates that non-checkers presented a higher probability of shift after losing than checkers in the negative-negative experiment.

Results from the two-way 2 x 2 ANOVA confirmed this interpretation. For win-shift values, it revealed that there was no significant main effect of experiment, group and no significant interaction between experiment and group. For lose-shift values, however, there was a marginal main effect of Experiment (Mean (Negative-neutral): 0.47±0.02; Mean (Negative-negative): 0.53±0.03; $F(1, 44) = 3.01, p = .09, \eta^2 = .06$), meaning that participants shifted less after losing in the negative-neutral experiment than in the negative-negative experiment (more exploration). In addition, there was a significant main effect of Group, (Mean (Checkers): 0.46±0.02; Mean (Non-checkers): 0.54±0.03; $F(1, 44) = 4.76, p = .04, \eta^2 = .10$), showing that checkers shifted less after losing than non-checkers. Finally, there was a significant interaction between the factors Group and Experiment, $F(1, 44) = 4.76, p = .04, \eta^2 = .10$. Post-hoc independent t-tests showed that checkers shifted less after losing in the previous trials than non-checkers in the negative-negative experiment (Mean (Checkers):
0.46 ± 0.04; Mean (Non-checkers): 0.61 ± 0.04; \( t(22) = 2.78, p_b = .02, r = .51 \), whereas no group difference was found for the negative-neutral experiment (Figure 6.10). Additionally, non-checkers shifted less in the negative-neutral experiment than in the negative-negative experiment (Mean (Negative-neutral): 0.47 ± 0.03; Mean (Negative-negative): 0.61 ± 0.04; \( t(20) = 2.71, p_b = .03, r = .50 \)). However, this difference between experiments was not evident for checkers in the negative study (Figure 6.10). In this way, non-checkers were influenced by the manipulation of the error magnitude in the negative feedback experiments by enhancing lose shifts and, hence, exploration strategies in the negative-negative experiment.

![Figure 6.10](image) Lose-shift mean values for checkers and non-checker for block 1 between the negative-neutral and negative-negative experiment in the negative study. Figure shows that checkers shifted less after losing than non-checkers in the negative-negative experiment and that non-checkers shifted less after losing in the negative-neutral experiment than in the negative-negative experiment.

In summary, win-shift values were not affected by reward magnitude manipulation in the negative direction. Lose-shift values did differ depending on the feedback magnitude and group. More specifically, non-checkers had enhanced lose-shift values in the negative-negative experiment compared to the negative-neutral experiment and compared to checkers, meaning that exploration behaviours increased with negative feedback magnitude.
enhancement in non-checkers. There was no experimental difference in lose-shift values for checkers.

6.3.2.2.5. Cross-correlation

6.3.2.2.5.1. Cross-correlation for Block 1

Figure 6.11 shows the auto- and cross-correlations for block 1 in the negative study, considering both groups and experiments together. Here, the difference between auto-and cross-correlation scores seems to be most evident for lag 1 and for lag 2. The three-way 5 x 2 x 2 mixed ANOVA on the auto- and cross-correlation difference scores revealed a significant main effect of lag, $F(4,166) = 15.33$, $p < .001$, $\eta^2 = .25$, as in Chapter 5. This was caused by a significantly higher difference between the participant’s auto- and cross-correlation values for lag 1 than for lags 2, 3, 4, and 5, all $t(47) > 3.50$, all $p$’s < .01. This might be because participants assumed a greater dependency between the current and the previous event (lag 1) than present in a random sequence. In addition, the difference between auto-and cross-correlations was higher at lag 2 compared to lags 3, 4 and 5, all $t(47) > 3.50$, all $p$’s < .01. This effect indicates that participants are not only using the previous trial result to make predictions as it was shown for previous studies but also using the second previous trial in both negative experiments (Figure 6.11). This effect was only found in this study. In this way, this indicates that in the negative experiments only, participants present a distinct behaviour concerning the use of the two previous trials to make assumptions about new decisions. No other significant finding was found.
Figure 6.11 Auto- and cross-correlations for lags 1 to 5 for block 1, while considering both groups and negative experiments together. The auto-correlation (averaged correlation values of the experimental sequences) is shown in black. Figure shows that participants presented a higher cross-correlation value than it was expected based on the experimental sequence for lags 1 and 2.

6.3.2.5.2. Cross-correlation for Block 2

Figure 6.12 shows that both groups generally followed the pattern in the experimental sequence in the negative study. In this way, checkers and non-checkers were able to learn sequential effects in this block. As expected the absolute cross-correlation values were not as large as the auto-correlation values. Figure 6.11 also reveals group differences regarding the participant responses in relation to the pattern sequence, where checkers performed better than non-checkers in the negative-neutral experiment, which was already reported in Chapter 5. However, there seems to be no group differences for the negative-negative experiment.
Figure 6.12 Auto- and cross-correlations for lags 1 to 5 for block 2. Comparison between checkers and non-checkers between the negative-neutral and the negative-negative experiments. The black curve represents the autocorrelation averaged across all event sequences. Figure shows a group difference in the negative-neutral experiment for lags 2 and 4.

This was reflected in the findings from the statistical analysis from the three-way 5 x 2 x 2 ANOVA for block 2. This analysis revealed a significant main effect of Lag, $F(2.12.,93.27) = 66.11$, $p < .001$, $\eta^2 = .60$ which was expected because of the pattern in the event sequence; a significant two-way interaction between the factors Lag and Group, $F(2.12.,93.27) = 3.12$, $p = .04$, $\eta^2 = .06$, and a significant three-way interaction between the factors Lag, Experiment, and Group, $F(2.12.,93.27) = 3.03$, $p < .050$, $\eta^2 = .06$.

The three-way interaction was further investigated by conducting a 2 x 2 between-participant ANOVA with the factors Group and Experiment. These ANOVAs revealed no effects for lags 1, 3, and 5. However, there was a significant interaction between Group and Experiment for lag 2, $F(1.,44) = 5.02$, $p = .03$, $\eta^2 = .10$ and a marginal interaction for lag 4, $F(1.,44) = 3.45$, $p = .07$, $\eta^2 = .07$.

This interaction at lag 2 was caused by significant group differences between checkers and non-checkers in the negative-neutral experiment, (Mean (Checkers): $-0.20 \pm 0.08$; Mean (Non-checkers): $0.57\pm 0.09$; $t(22) = 2.89$, $p_{b} = .02$, $r = 0.52$), where checkers were closer to
the auto-correlation values than non-checkers in lag 2 which means that they learned the pattern better (Figure 6.12). These findings were already shown in Chapter 5. More interestingly, no group difference was found for the negative-negative experiment.

For lag 4, again, this was caused by group differences between checkers and non-checkers in the negative-neutral experiment (Mean (Checkers): 0.13 ± 0.06; Mean (Non-checkers): 0.39± 0.07; \( t(22) = 2.60, p_b = .03, r = 0.48 \)), where checkers were closer to the auto-correlation value at lag 4 than non-checkers which means that they learned the sequence better (Figure 6.12). More interestingly, this group difference was not present in the negative-negative experiment.

6.3.2.2.5.3. Cross-correlation for Block 3

The three-way ANOVA 5 x 2 x 2 for block 3 revealed a significant main effect of Lag, \( F(1.79.,78.93) = 49.22, p < .001, \eta^2 = .52, \) that was expected because of the presence of the pattern sequence, and a marginally significant interaction between the factors Lag and Group, \( F(1.79.,78.93) = 2.79, p = .07, \eta^2 = .06. \) The latter interaction was further investigated with five independent post-hoc t-tests comparing groups (with both experiments averaged together) for each lag separately. Results have shown that there was a marginally significant group difference for lag 1 after a Bonferroni correction (Mean (Checkers): -0.14 ± 0.04; Mean (Non-checkers): -0.27± 0.04; \( t(46)= 2.33, p_b = .1, r = .33 \)), indicating that checkers were numerically closer to the experimental sequence (auto-correlation) than non-checkers, meaning they might take the immediately trial more into account when making their decisions about the current trial than checkers, when considering both negative experiments together (Figure 6.13).
Figure 6.13 Auto- and cross-correlation values for lags 1 to 5 of block 3. Comparison between checkers and non-checkers considering both experiments together for the negative study. The black curve represents the averaged auto-correlation for the experimental sequence. Figure does not indicate any group difference when considering both experiments together.

In summary, the findings from the auto- and cross-correlation analysis indicate that the two preceding trials influenced the decision making in block 1 in experiments with negative feedback. For block 2, checkers cross-correlations were closer to the auto-correlation values than non-checkers cross-correlation values when assessing lags 2 and 4 of the negative-neutral experiment, which reflects better pattern learning in checkers. This was already reported in Chapter 5. More interestingly, no group differences were found for block 2 in the negative-negative experiment. For block 3, no significant group or experimental effects were found. These findings will be discussed in Chapter 6.4.

6.3.2.3. Questionnaires

6.3.2.3.1. Intolerance of uncertainty scale (IUS)

The two-way 2 x 2 ANOVA with the factors Group and Experiment revealed a marginally significant main effect of Group for factor 1 (Uncertainty has negative behavioural and self-referent implications); Mean (checkers): 37.54 ± 2.12; Mean (non-checkers): 31.71± 2.33; $F(1,44) = 3.39$, $p = .07$, $\eta^2 = .07$), a significant main effect of Group
for Factor 2 (uncertainty is unfair and spoils everything); Mean (checkers): 35.96 ± 1.94; Mean (non-checkers): 28.21 ± 1.89; $F(1,44) = 7.91$, $p = .007$, $\eta^2 = .15$), and a significant main effect of Group for the total IUS score (Mean (checkers): 73.05 ± 3.86; Mean (non-checkers): 59.92 ± 4.15; $F(1,44) = 5.62$, $p = .02$, $\eta^2 = .11$). This shows that checkers had higher intolerance of uncertainty ratings than non-checkers. This supports similar findings from Chapter 5 and Chapter 6 (Positive study).

6.3.2.3.2. Domain Specific Risk-Taking Scale and Risk-Perception scale (DOSPERT)

The two-way 2 x 2 ANOVA with the factors Group and Experiment showed no significant main effects and interaction for finance risk-taking and risk-perception scores or the total risk-taking and risk-perception scores. Thus, it seems checkers and non-checkers rated their risk taking and risk-perception behaviour similarly.

6.3.2.3.3. UPPS Impulsivity Scale

The two-way 2 x 2 ANOVA did not have significant main effects or interaction effects for any of the UPPS sub-sections and total score. Thus, it seems checkers and non-checkers rated their impulsivity similarly.

6.3.2.3.4. South Oaks Gambling Screen

The two-way 2 x 2 ANOVA revealed a marginally significant main effect of Group for the gambling ratings (Mean (checkers): 0.75 ± 0.20; Mean (non-checkers): 0.29± 0.13; $F(1,44) = 3.60$, $p = .06$, $\eta^2 = .07$) showing that checkers had numerically enhanced levels of gambling behaviour compared to non-checkers.

6.3.2.3.5. Strategy questionnaires

Categorical data from the strategy questionnaire showed that checkers and non-checkers seem to employ strategies and to visualize patterns in the same way in the negative-negative experiment for all blocks. Firstly, there were no strategy differences between groups and experiments for block 1. For block 2, checkers detected marginally significantly more often the pattern of block 2 in the negative-neutral experiment (Checkers: 11 yes, 1 no; Non-checkers: 6 yes, 6 no; $p = .07$) which was already reported in Chapter 5. There was no
difference between experiments for the checker and non-checker group for block 2. For block 3, non-checkers reported more pattern detections in the negative-negative experiment (11 yes, 0 no) than in the negative-neutral experiment (5 yes, 7 no; \( p = .005 \)), whereas the number of pattern detections was similar between experiments for checkers in the negative-neutral (8 yes, 4 no) and in the negative-negative experiment (10 yes, 2 no), \( p = .64 \).

In summary, the analysis of the lab-based questionnaires revealed that checkers have higher intolerance of uncertainty ratings. Checkers also gambled numerically more than non-checkers but this effect was only marginally significant. In addition, the strategy questionnaire revealed no group or experimental differences for block 1. However, checkers reported the detection of a pattern in block 2 more often than non-checkers in the negative-neutral experiment compared to the negative-negative experiment. There was no difference between experiments for non-checkers. In block 3, non-checkers detected a pattern less often in the negative-neutral experiment than in the negative-negative experiment. There was no difference between experiments for the checkers.

6.3.2.4. Correlation analysis between behavioural measures from the binary decision making task and laboratory-based questionnaire scores

6.3.1.4.1. Correlation analysis for the negative-neutral experiment

This analysis was already reported in Chapter 5. As a reminder, the Kendall Tau correlation analysis on checkers revealed the following findings. Firstly, symptom severity (checking OCI-R) was negatively correlated with performance as reflected in cross-correlation values in block 2 (Diff lag 4: \( r_\tau = .46, p = .04 \)). Secondly, comorbidity levels did not correlate with the binary decision making task performance. Finally, laboratory-based questionnaire ratings correlate with accuracy in block 2 in the binary decision making task (Factor 1: \( r_\tau = -.61, p = .006 \); Factor 2: \( r_\tau = -.47, p = .03 \); Total score: \( r_\tau = -.54, p = .016 \)) and pattern learning in lag 2 (Factor 1: \( r_\tau = -.54, p = .01 \); Factor 2: \( r_\tau = -.46, p = .03 \); Total score: \( r_\tau = -.46, p = .03 \)). In this way, higher levels of intolerance of uncertainty were related to a decreased performance in the binary decision making task for block 2. In this experiment, intolerance of uncertainty is influencing behaviour, what is the opposite of what was found for the positive experiments (Positive study).
Correlation analysis for the negative-negative experiment

For the negative-negative experiment, symptom severity (OCI-R total) was negatively correlated with the win-shift measure, $r_t = -.45, p = .04$. In addition, the OCI-R total was also negatively correlated with the auto- vs. cross-correlation difference at lag 1 for block 1, $r_t = -.47, p = .03$, which indicated that checkers with higher OCI-R total scores were less likely to shift after winning (more exploitation) and would rely less on the previous trial information to make their decisions. Secondly, higher comorbidity levels did not correlate with the binary decision making task performance. Thirdly, laboratory-based questionnaire ratings did correlate with the binary decision making task performance. For block 1, intolerance of uncertainty negatively correlated with the lose-shift measure of the binary decision making task (IOU factor 2: $r_t = -.47, p = .03$, IOU total: $r_t = -.47, p = .03$). These negative correlations indicate that higher intolerance of uncertainty in checkers was related to less shifts after losing and, hence, to enhanced exploitation behaviour. For block 2, intolerance of uncertainty ratings positively correlated with accuracy levels in block 2, $r_t = .45, p = .04$. In addition, intolerance of uncertainty ratings (factor 2) negatively correlated with the difference between the experimental sequence auto-correlation and the participants cross-correlation values in lag 2 for block 2, $r_t = -.46, p = .03$, meaning that more intolerance of uncertainty in the IUO rating (factor 2) increases pattern learning in block 2. This result is opposite to the finding for the negative-neutral experiment, where checkers with higher values of intolerance of uncertainty were the ones with the lowest accuracy in block 2 (see section 6.3.1.4.1. and Chapter 5).

In summary, checkers with higher symptom severity learned less in block 2 in the negative-neutral experiment (as already reported in Chapter 5). This effect was not present for the negative-negative experiment. However, enhanced symptom severity was related to reduced win-shift values and less usage of the information from the previous trial (lag 1) in block 1 of the negative-negative experiment. This reflects a behaviour that is less biased towards exploration in checkers with enhanced symptoms. This correlation was not found for the negative-neutral experiment. When looking at the relationship between intolerance of uncertainty and binary decision making task performance, a similar picture occurs for the negative neutral experiment. Here, enhanced IOU levels are related to reduced accuracy levels in block 2. Interestingly, this effect was opposite for the negative-negative experiment where enhanced IOU levels were positively related to enhanced accuracy levels in block 2. Additionally, IUO also negatively correlated with the lose-shift measure of block 1 in the
negative-negative experiment, reflecting more exploitation behaviour with higher intolerance of uncertainty. Note that while in the negative study, intolerance of uncertainty seems to be influencing decision making in checkers, intolerance of uncertainty seems to not affect decision making in checkers in the positive study. These correlational findings will be discussed together with findings from previous studies in Chapter 8.

6.4. Discussion

OCD individuals have deficits in decision making tasks under ambiguity which are often equivalent to probabilistic reward-based learning tasks (Starcke et al., 2010; Kim et al., 2015; Zhang et al., 2015). However, research findings are contradictory and might be due to task differences. For example, tasks as the IGT simultaneously present distinct feedback directions and feedback magnitudes which makes it difficult to know which variable is deficient in OCD. This present thesis tried to disentangle these factors and examine the influence of distinct feedback directions (Chapter 5) and feedback magnitudes in the task (Chapter 6).

Here, the present two studies investigated subclinical checkers and matched non-checkers and compared their performance while manipulating values of feedback magnitude. The positive study manipulated the feedback magnitude between two experiments within a positive feedback direction and the other study manipulated the feedback magnitude between two experiments within a negative feedback direction. More specifically, feedback magnitude effects were compared between the positive-neutral experiment (positive monetary and neutral picture feedback vs. absence of monetary and neutral picture feedback) and the positive-positive experiment (positive monetary and positive picture feedback vs. absence of monetary and neutral picture feedback) in the Positive study. In the negative study, feedback magnitude effects were compared between the negative-neutral experiment (negative monetary and neutral picture feedback vs. absence of monetary and neutral picture feedback) and the negative-negative experiment (negative monetary and negative picture feedback vs. absence of monetary and neutral picture feedback). Positive and negative experiments were analysed separately and are reported in two distinct studies because the effect of feedback magnitude could be rather different for these two response directions. Note that the separate comparisons between the two positive experiments and the two negative experiments with distinct feedback magnitudes made it possible to compare the effects of the presence vs. absence of reward magnitudes and presence vs. absence of punishment.
magnitudes separately. Although the analysis of the positive and negative experiments was separated – the discussion will integrate the findings.

In addition to the feedback manipulation, the event sequence type was altered from block to block in all experiments in order to investigate whether checkers are better / worth learners in random sequence RBL or in pattern sequence RBL. Hence, three different blocks were presented in all experiments, i.e. a random sequence block with a target-option manipulation, a pattern sequence block with equal winning probabilities for both target options, and a pattern sequence block where one target option had a higher chance of winning.

Similarly to previous findings from Chapter 5, the findings from the accuracy analysis of all three blocks showed that participants were able to learn the task over time. Block 1 was more difficult than blocks 2 and 3 (lower maximum accuracy), as participants needed two sub-blocks to increase accuracy, instead of just one in blocks 2 and 3. No group accuracy differences were found for the random sequence block with differential target option probabilities (block1), supporting results from Chapter 5. There was one exception, a marginally significant effect of experiment in the positive experiment analysis which might suggest that a feedback magnitude enhancement in the positive direction might also numerically enhance the performance in the random-sequence blocks, i.e. the accuracy was numerically higher for the positive-positive compared to the positive-neutral experiment.

6.4.1. Discussion of the feedback magnitude effect on RBL with random-sequences in checkers and non-checkers

The first aim of this study was to compare the effects of feedback magnitude in a random sequence RBL block between people with subclinical checking symptoms and non-checkers. The findings from the accuracy analysis for block 1 did show that the accuracy was numerically higher for the positive-positive compared to the positive-neutral experiment. No group differences were found for the negative feedback experiments. In addition to the accuracy analysis in the random sequence block, the usage of exploration /exploitation strategies was evaluated. Findings showed exploration / exploitation differences between the subclinical checker group and the non-checker group in both studies.

For the positive experiments, the findings from the win-shift analysis were in accordance with the results from the accuracy analysis, as there was a marginal effect of experiment for win-shift, where participants shifted less between options in the positive-
positive experiment than in the positive-neutral experiment, which means they were closer to the optimal strategy of sticking to one option (exploitation) after winning, when the feedback magnitude increases. This could also explain the numerical accuracy differences between the positive feedback experiments described above.

For the lose-shift probabilities, however, a group difference occurred. Only checkers shifted less after losing in the positive-positive experiment than in the positive-neutral experiment, which reflects a disease-related enhanced exploitation effect in subclinical checkers. This result is in line with the cross-correlation findings, where checkers were less influenced by the immediately preceding trial (lag 1) in the positive-positive experiment when compared to non-checkers. In this way, they were biased towards ignoring the feedback about the previous decision, especially after the absence of positive feedback. The absence of positive feedback in the positive-positive experiment seems to trigger the exploitation behaviour in checkers.

This finding indicates that checkers and non-checkers were similarly affected by winning (positive feedback present) because they enhanced their exploitation behaviour after winning with enhanced feedback magnitude in the positive experiments. However, checkers were more influenced by the absence of positive feedback and this influence could be measured with strategy changes towards exploitation and cross-correlation changes in lag 1. In summary, it was predicted that checkers are specifically influenced by absent feedback (which is perceived as negative feedback) in the positive-positive experiment.

Interestingly, the group differences in the lose-shift and cross-correlation measures were not reflected in accuracy differences between groups. This is potentially because these measures are more sensitive to changes than the accuracy measure. This shows that additional strategy measures might be able to show effects that classical accuracy levels cannot show.

For the negative experiments, the analysis revealed, firstly, no group differences between checkers and non-checkers when comparing them in the negative-neutral experiment. Secondly, there were also no accuracy and win-shift probability differences between groups for the negative-negative experiment, although the feedback magnitude was enhanced in the negative-negative experiment. Interestingly, however non-checkers shifted more after losing in block 1 (more exploration) in the negative-negative experiment. This means that non-checkers explore more the negative environment after losing in order to find another alternative strategy, whereas checkers do not adapt their strategy behaviour. Based on the previous literature (Simon et al., 2010), it was hypothesized that checkers present a
bias towards exploitation when exposed to aversive stimuli in experiments with a negative feedback as OCD individuals are oversensitive to these stimuli. This was indeed the case. Interestingly, this strategy difference occurred because checkers did not adjust to the challenges of the environment whereas non-checkers did adjust by becoming explorers.

Unfortunately, the cross-correlation findings from the negative experiments did not directly support the lose-shift findings because no group differences were found for this measure in the negative experiments. However, there was another interesting finding: participants were being influenced by the two preceding trials to make decisions. This means they were more influenced by preceding information than in the positive experiments where only a lag 1 effect was found. This could suggest more pattern search behaviour in the negative experiments than in the positive experiments which supports the assumption about pattern search in random event sequences (see Vulkan, 2000). This enhanced pattern search does fit well to the enhanced exploration behaviour in non-checkers (lose-shift measure) but not to the checker behaviour, so these results are possibly reflecting non-checker exploration.

It was argued in the general introduction that the presence of a penalty is not the same as the absence of a reward, as they seem to be related to distinct physiological systems (Schultz, 2010). Our findings support this differentiation between the feedback forms. More specifically, one could consider win-shift probabilities as a reflection of the effect of positive feedback (positive experiments) or the absence of negative feedback (negative experiments) on decision making in the next trial, while lose-shift probabilities are a measurement of the effect of errors on decision making. Here, these errors are either linked to the absence of reward (positive experiments) or the presence of penalties (negative experiment). In other words, lose-shift deficits could be indirectly indicating deficits in negative prediction errors. Prediction errors are the basis for reinforcement learning because they conceptualise the comparison between the received unexpected feedback and the internal model about what should have happened in a trial (Maia, 2009). Based on this comparison the next action will be decided upon – e.g. shift or stay on same option. Note, however that win-shifts could not be indirectly linked to positive prediction errors, because positive prediction errors are released after unexpected winnings, for example, when one receives two rewards, instead of one reward that was already expected. Our studies did not manipulate distinct amounts of rewards in the same task, so there were no unexpected winnings, what is different for errors that are usually unexpected.
Interestingly, in all four experiments no group differences between checkers and non-checkers were found for win-shift probabilities which changed equally for both groups. This suggests that checkers do not have any deficit in processing and applying positive feedback. However, findings suggest that checkers have a deficit when negative feedback is given. There were two effects: Firstly, the absence of reward in the positive-positive experiment resulted in higher exploitation behaviour in checkers compared to the absence of reward in the positive-neutral experiment. Note that checkers adjust their behaviour because of feedback changes in a positive learning environment. In contrast, the presence of higher levels of punishment (indirectly linked to higher negative prediction errors) in the negative-negative experiment compared to the negative-neutral experiment enhanced lose-shift probabilities in non-checkers compared to checkers. Here, non-checkers adjusted their behaviour to the negative environment and started to explore their options. Checkers did not perform this adjustment. This could be because checkers are not able to adjust to negative prediction errors of certain magnitudes and that they are unable to correct their strategies based on previous feedback.

Summarising the findings from the positive and negative studies, they indicate that checkers do present deficits when exposed to enhanced negative feedback that surpasses a threshold value. However, the specific form of deficit depends on the negative feedback type in relation to environmental factors. When receiving an enhanced negative feedback (absence of positive feedback) magnitude in a positive feedback environment, checkers adjust their behaviour to become more exploiters. When receiving an enhanced negative feedback (presence of penalties) in a negative feedback environment, they do not adjust their behaviour towards exploration and remain exploiters.

These findings are supported by animal studies that investigated reward and punishment in RBL while looking at neurotransmitter systems. According to Schultz (2010), the absence of a reward is not modulated by the same physiological system as the presence of a penalty. Schultz (2010) reported that the absence of a reward is not followed by a negative prediction error, at least, not one that is controlled by the dopamine system. Indeed, while dopamine systems seem to generate positive prediction errors and maybe negative prediction errors (see Maia, 2009 for a discussion), not much is known about the effect of the absence of reward in RBL (Schultz, 2010). Actually, Fiorillo (2013) argued that others neurotransmitter systems, as the serotonergic system, could be responsible for prediction errors related with the absence of a reward, or the absence of punishment.
Another potential explanation for our findings, which is not diminishing the previous explanation, could be that there are distinct systems modulating exploration and exploitation in the brain and the exploitation system is overactive in OCD, which could explain the deficits in the subclinical checkers. According to the study results for non-checkers, the correct decision for healthy individuals would be to do more exploration when the value of an error significantly increases in negative environments. In other words, when the value of losing is lower (in terms of monetary and emotional values), as in the case of the positive-neutral and negative-neutral experiments, in relation to the positive-positive and negative-negative experiments, healthy individuals could risk more by exploring more strategic options while disregarding actual results that could be more related with noise than to real information. Actually, participants are never sure if negative feedback discredit their predictions about the optimal strategy or if this feedback is just due to the noise level in the task (see Sutton & Barto, 1999 for more details). Therefore, it is best for participants to balance exploration with exploitation behaviour in a RBLT (Sutton & Barto, 1999).

However, if the risk of losing is lower or the loss presents a lower value, participants might risk more and consequently exploit their predictions based on their internal model about the learning environment. However, when the risk of losing or the value of the loss increases, as in the negative-negative experiment, participants might consider the negative feedback more carefully and explore the environment for new strategies, while trusting less their internal model. Interestingly, activations of frontopolar cortex and intraparietal sulcus are associated with exploration behaviour (Daw et al., 2006), whereas activations of ventral and dorsal striatum, as well as dopaminergic networks, are more linked to exploitation strategies (Montague et al., 1996; Delgado et al., 2000; Daw et al., 2006). According to the authors, the fact that exploration is linked to the frontopolar cortex, an area implicated in behavioural control (Miller & Cohen, 2001), could support a theory that exploration is accomplished by overriding exploitative tendencies when necessary (Kaelbling, 1993). It is possible that an overactive orbitofrontal-striatal system (Menzies et al., 2008), which is innervated by a hyperactive dopamine system in OCD (e.g. Perani et al., 2008), could bias participants with obsessive tendencies to use exploitation strategies, even when explorative tendencies are needed for a better task performance.

As a last point, correlations between the binary decision making task performance in the random-sequence block and the questionnaires were performed to investigate relationships between the performance in the binary decision making tasks and OCD.
symptoms, potential OCD comorbidities such as anxiety and depression, and other potential factors such as risk taking, impulsivity, intolerance of uncertainty, and gambling behaviours. Please note, these correlations were only performed on very small participant numbers. They were not similar between experiments and it is highly possible that these results are related with specific sample characteristics and not the factors described above. Hence, effects should be treated with caution and need further investigation in future studies.

Nevertheless, a few interesting correlation in relation to block 1 will be mentioned here. In fact, anxiety was related to enhanced lose-shift probabilities in the positive-neutral experiment, meaning that they become more explorers after losing in mildly positive environments. Note, there was no correlation between anxiety and performance in the positive-positive experiment where rewards were higher but the absence of reward was also more negatively felt. Interestingly, this finding also shows that anxiety might have different effects on the performance in this task. This finding might be worth exploring in future studies.

In summary, findings from the positive and negative experiment indicate that checkers do present deficits when exposed to enhanced negative feedback. The specific form of deficit depends on the negative feedback type in relation to environmental factors. When receiving absent positive feedback in a positive feedback environment, checkers adjust their behaviour to become more exploiters. When receiving negative feedback in a negative feedback environment, they do not adjust their behaviour towards exploration and remain exploiters. These findings could be explained with models that either assume that the four feedback types used in this task are linked to different neurotransmitter systems (Fiorillo, 2013) or by models that assume two distinct exploration and exploitation brain systems (Daw et al., 2006).

### 6.4.2. Reward-based learning of pattern sequences in uncertain environments

A second aim was to extend the investigation of probabilistic reward-based learning tasks from random sequences to pattern sequences when considering the manipulation of feedback magnitude. In probabilistic environments, decision making in RBLTs with pattern sequences might be different to decision making in RBLTs with random-sequences (see Chapter 1 for a detailed discussion of this issue). Moreover, checking-related tendencies might affect decision making in RBLTs with random-sequences differently than decision making in RBLTs with pattern sequences. Findings from Chapter 5 support this suggestion.
They showed that group differences between checkers and non-checkers performing in a positive-neutral or negative-neutral experiment were specifically prominent for the pattern sequence block 2 of the negative-neutral experiment. More specifically, checkers performed better in this block compared to non-checkers which is likely to be linked to higher exploitation strategies in checkers. In this way, negative feedback magnitude is not the only factor that seems to affect decision making under ambiguity / RBL in checkers. Differences in optimal RBL strategies between blocks need to be considered to.

For the pattern sequence blocks 2 and 3, it was expected that checkers would present a bias towards exploitation in the positive-positive and negative-negative experiments in comparison with the positive-neutral and negative-neutral experiments, based on the hypothesis that checkers are affected by a larger absence of reward in the positive-positive experiment or by a larger presence of punishment. This bias would result in better accuracy levels and better pattern learning cross-correlations for checkers compared to non-checkers with enhanced feedback magnitudes.

For the positive experiments, findings from blocks 2 and 3 did not show any group or experimental accuracy and differences for both positive experiments. There was only a significant cross-correlation difference between groups for lag 3 showing that non-checkers were closer to the experimental sequence in lag 3 than checkers in block 2, considering both experiments. This effect was rather odd because pattern learning effects should have been reflected in lag 2 and lag 4 differences. However, this corroborates findings from Chapter 5, where non-checkers were closer to the pattern sequence than checkers in the positive-neutral experiment in lag 5. Actually, this would indicate that non-checkers are more biased towards exploitation in the positive experiments. However, these findings should be treated with some caution, as no significant findings were found for lags 2 and 4. New studies should be designed to replicate these findings with a bigger sample before giving it a bigger meaning. Interestingly, the strategy questionnaire for the pattern in block 2 showed that checkers have detected more patterns in the positive-positive experiment than in the positive-neutral experiment. This could indicate that checkers are doing more exploitation in this block and hence, seeing more patterns here, what would corroborate findings for block 1, where checkers did more exploitation in the positive-positive experiment. However, a limitation of the structure questionnaire is that we do not know which pattern was seen or imagined. There was no verification whether the correct pattern was detected. Additionally, these results were not corroborated by behavioural results related with accuracy or cross-correlations. Hence,
the strategy questionnaire results cannot be well interpreted. For block 3, there were no group or experimental effects for the cross-correlation measures, as in Chapter 5. In addition, there was no significant group differences for the strategy questionnaire. Strategy detection was more difficult in this block because two strategies were intermixed here, so it is possible that the presence of two optimal strategies in the same block could be responsible for the absence of effects regarding block 3.

The findings from the negative experiments showed, firstly, that checkers had higher accuracy levels and were closer to the pattern sequence than non-checkers in block 2 of the negative-neutral experiment. In accordance, checkers reported the detection of a pattern in block 2 more often than non-checkers in the negative-neutral experiment compared to the negative-negative experiment. There was no difference between experiments for non-checkers. This analysis was already reported and discussed in Chapter 5. Secondly, and more interesting, this group difference was not present in the negative-negative experiment. There are several explanations for this effect. Actually, checkers did not perform differently between the negative-neutral and negative-negative experiment and, hence, their exploitation levels are similar in both experiments. This would be in line with the findings from the random-sequence block analysis (block 1). Note, however, there was also no experimental difference for the non-checkers. A second explanation could be that the absence of experimental differences could be due to low sample size issues. Finally, it is possible that results of a bias toward exploitation in block 2 of the negative-neutral experiment was due to between group sampling errors. In this way, it would be interesting to replicate the negative-neutral experiment to confirm the exploitation bias in checkers.

For completeness, no group or experimental differences were found for block 3 when analysing the negative experiments, confirming the findings from Chapter 5. This might be due to the mix of strategies that can blur the effects. Note, the strategy questionnaire showed that non-checkers detected more patterns in block 3 of the negative-negative experiment than in block 3 of the negative-neutral experiment. However, this was not reflected in the performance data.

An explanation for the absence of group effects in the pattern blocks, especially of the positive-positive and negative-negative experiments, could be that non-checkers are generally biased towards exploitation, when the maximum accuracy is rather high, as in blocks 2 and 3, irrespectively of the fact the error magnitude of the experiment is high. This would decrease group differences in these conditions. This would be the case because higher values of
maximum accuracy would decrease the error probability in the task and, consequently, decreasing the risk of opting for exploitation in non-checkers. Note that checkers might not be able to adapt and would always opt for exploitation. However, future studies should be designed to investigate the effects of feedback probability manipulation in the exploitation / exploration bias of subclinical checkers.

As the last point, for the negative experiment, the correlational studies showed that symptom-severity and intolerance of uncertainty ratings were related to changes in binary decision making task measures. No correlations were significant for the comorbidities or for the other laboratory based questionnaires. When looking at the relationship between symptom-severity and task measures the following findings are interesting. In the negative-neutral experiment, symptom severity was negatively correlated with the accuracy levels in block 2 (as already reported in Chapter 5). This effect was not present for the negative-negative experiment and, hence, the initial finding should be treated with more caution before being replicated. When looking at the relationship between intolerance of uncertainty and binary decision making task performance, a similar picture did occur for the negative-neutral experiment. Here, enhanced IOU levels were related to reduced accuracy levels in block 2. Interestingly, this effect was opposite for the negative-negative experiment where enhanced IOU levels were positively related to enhanced accuracy levels in block 2. This is a rather unusual finding that needs to be replicated. Currently, we cannot give an explanation for this. Note while in the negative study, intolerance of uncertainty seems to be influencing decision making in checkers, intolerance of uncertainty seems to not affect decision making in checkers in the positive study.

In summary, results from the positive study and the negative study partially supported the hypothesis that checkers might present a specific deficit regarding higher values of negative feedback magnitudes. These values seem to be above a specific threshold. Indeed, these deficits seem to not be related with feedback direction but with feedback magnitude because deficits were both found in the positive and in the negative feedback direction. Additionally, it seems that when values of error magnitudes surpass a specific value of negative feedback magnitude, checkers are unable to adapt behaviour and present a bias towards exploitation. This might indicate that checkers are unable to override exploitative tendencies in the brain, when more exploration is required. It is possible this is the case because the systems associated with exploitation are overactive in OCD (Menzies et al., 2008), as the dopaminergic and striatal network (Daw et al., 2006).
could lead to deficits in negative prediction errors. Besides the influence of the error magnitude, it is possible that this bias is also affected by the probability associated with the targets that are reflected by values of maximum accuracy.

6.4.3. Limitations and future directions

There are several potential limitations of the positive and negative study. Firstly, it would be important to replicate the findings from this study, considering a larger sample size of sub-clinical checkers. Secondly, the use of a subclinical checker sample with a lower symptom severity could be another reason for negative, non-significant, findings. Therefore, it would be useful to replicate this study with a clinical sample. Thirdly, no direct physiological measures of emotional sensitivity to negative and positive magnitudes was taken, e.g. skin conductance responses (SCR) and heart rates, and it would be quite interesting to understand how the manipulations of feedback magnitudes are sensed by participants within the positive and the negative direction. Indeed, it would be also possible to investigate if the absence of a reward is similarly felt as the presence of a penalty in experiments with similar amounts of rewards and penalties. Additionally, fMRI studies could be designed to correlate behavioural measures with exploration and exploitation systems in the brain. The role of an overactive dopamine system could be studied with participants screened for genetic phenotypes. Finally, these RBLTs could also be used to investigate decision making deficits in Parkinson’s disease patients (on and off medication) to understand the role of dopamine changes in these tasks.

6.5. Conclusions

Studies 1 and 2 corroborate findings of Chapter 5. Actually, decision making in checkers and non-checkers seem to vary, depending on the feedback direction and magnitude of the task and on the optimal strategy associated with the task. Here, checkers seemed to shift less in conditions where the error magnitude was higher. While results corroborate the hypothesis of a deficit regarding highly negative error magnitudes in checkers, some results that were expected were not found, particularly for the blocks with patterns. In this way, this conclusion needs to be taken carefully. However, it is still important to understand how checkers are affected by distinct error magnitudes. Especially, there are no known studies that tried to investigate how OCD decision making is affected by symptom-related reinforcers. Until now, it is possible to know that OCD participants are highly affected by symptom-
related emotional stimuli as by general aversive stimuli (see Simon et al., 2010) but how does this affect decision making, in comparison with a general negative stimuli? This question will be addressed in Chapter 7.
Chapter 7 - Reward based-learning in subclinical checkers: Symptom-specific alteration of reward magnitude

7.1. Introduction

As seen before, OCD population is characterized by a specific decision making deficit, regarding reward-based learning tasks, also known as decision making under ambiguity (Starcke et al., 2010). Tasks that are usually employed to investigate these deficits present distinct factors at the same time, as feedback direction and feedback magnitude, so previous Chapters tried to disentangle feedback direction (Chapter 5) and feedback magnitude (Chapter 6) to separately examine the effect of the manipulation of these factors in the decision making of OCD. Additionally, studies from these chapters also examined the influence of sequential effects and first order effects (use of the immediately previous trials), using blocks with sequences presenting random sequences without patterns (block 1) and sequences with patterns (block 2 and 3).

These previous studies support that subclinical checkers might present a deficit towards specific values of error magnitudes. In this case, it was shown that deficits could be present in the positive and in the negative feedback direction but when the magnitude of the error, related with the task is augmented, subclinical checkers seem to be biased towards exploitation and decrease exploration in the task. Interestingly, besides being biased towards exploitation, these studies indicate that subclinical checkers are not able to adapt when the error magnitude of the task increases. In this way, while non-checkers seem to change the bias towards exploration, according to the error magnitude of the negative task, checkers always present similar amounts of exploitation in the task. In this way, this could indicate that subclinical checkers are not able to adapt their strategies, when the error magnitude of the task surpasses a certain value but it is still not possible to know how a symptom-related aversive stimuli would affect their decisions and if it would influence subclinical checkers to do even more exploitation in the task.

In fact, from my knowledge, no RBL studies concentrated in the investigation of the influence of symptom-related stimuli in reward-based learning in OCD. This is especially important because symptom-related stimuli usually enhances fronto-striatal activation in OCD and this could reflect an emotional hyperarousal to these stimuli (Simon et al., 2010), what could consequently trigger decision making deficits or more exploitation. In this way, it is possible that tasks presenting symptom-related aversive feedback could trigger specific
deficits in OCD or decrease performance even more than that what was previously reported with general aversive stimuli.

In fact, it is still not known if these deficits will be stronger when subclinical checkers are exposed to symptom-related aversive stimuli. In this way, to pursue the investigation of RBL deficits in OCD, one new study was designed to investigate the influence of these specific negative symptom-related emotional stimuli in the RBL of subclinical checkers and if those stimuli differently affect checker behaviour due to its nature. In this way, subclinical checker results were compared between the negative-neutral experiment (study 3) and a symptom-related experiment (study 6). In the negative-neutral experiment participants received some money to gamble and they did not lose anything after finding the target (absence of penalty); meanwhile, if participants made a mistake, they lost money and saw a neutral picture (presence of penalty). In the symptom-related experiment, participants received some money to gamble and they did not lose anything after finding the target (absence of penalty); meanwhile, if participants made a mistake, they lost money and saw an aversive symptom-related picture (presence of penalty). Results were compared between checkers and non-checkers for each experiment and for each group between experiments.

Actually, this design was similar to the one of the Negative study (negative-neutral vs. negative-negative), as both experiments presented the same feedback direction but distinct feedback magnitudes. In fact, if feedback magnitude value is influencing the bias towards exploitation in subclinical checkers but not the nature of the stimuli per se, the symptom-related stimuli will not differently affect decisions in checkers. Indeed, the negative valence associated with the symptom-related pictures are higher than the neutral pictures delivered at the negative-neutral experiment, so it would be expected that results would be similar to the ones of the Negative study in Chapter 6 (negative-neutral vs. negative-negative), where checkers would present a bias towards exploitation in comparison to non-checkers and would not present any difference between experiments. In addition, it is possible that non-checkers will present different results associated with each experiment, as it was shown in the Negative study (see Chapter 6) with more exploration being performed in the experiment with higher error magnitude (symptom-related experiment as study 6).

However, if symptom-related stimuli distinctively affects checkers, not due to its feedback magnitude but because of its nature regarding the symptom-related obsessions associated with it, it will be expected that checkers will present more exploitation in the symptom-related experiment than in the negative-neutral experiment, so a difference between
experiments will be found for checkers. Based on results from the Negative study in Chapter 6, it is possible that these differences will be reflected in block 1 but not in blocks 2 and 3.

7.2. Methods

7.2.1. Overview of the study design

This study design was exactly the same as the ones shown at Chapters 5 and 6. In this way, students completed an online screening questionnaire, selected participants were randomly sorted to perform the symptom-related experiment (study 6), they completed the picture validation task, performed the binary decision making task and finally they completed the lab-based questionnaires.

7.2.2. Participant selection based on the online screening questionnaire

The Online screening questionnaire consisted of some demographic questions about age, gender, play videogame and play for money, past or present diagnosis of depression, anxiety, panic disorder and post-traumatic stress disorder, the Obsessive Compulsive Inventory revised (OCI-R) and the Depression, anxiety and stress scale (DASS) questionnaires (Appendix 2.3). Methods about the online questionnaire can be seen in detail in Chapter 2.

Participant selection was done together for the positive-positive (study 4), negative-negative (study 5) and symptom-related experiment (study 6). In this way, details about exclusion can be seen at table 6.1 at Chapter 6.2.2. Basically, from 352 people who signed in for the study, 24 participants complete the symptom-related experiment. In each experiment, 12 participants were checkers and 12 were non-checkers.

7.2.2.1. Participants on the laboratory-based study: Group sample characteristics

For tests of normality, z-scores for skewness and kurtosis were calculated for age, and the DASS and OCI-R scores. Gender, videogame experience, habit to play for money were investigated as categorical variables. DASS depression was not considered as a normal variable. For the OCI-R questionnaire, OCI-R hoarding in the symptom-related was not considered normal.

Online questionnaire variables were compared across groups and experiments to compare the selected group samples. For most variables, this comparison was conducted by using between-participant ANOVAs with the factors Group (checker vs. non-checker) and
experiment (negative-neutral vs. symptom-related). Gender, videogame playing and gambling with money were the exceptions because these were categorical variables. Here, Chi-Square tests or Fisher’s exact tests were used to explore sample differences between groups or experiments. The findings from this analysis are presented in Table 7.1.

For both the OCI-R and the DASS scores, ANOVAs did reveal main effects of Group, except for DASS depression (Table 7.1). Basically, checkers always had higher scores than non-checkers in all OCI-R and DASS subscales, independent of the experiments (Table 7.1). For the DASS depression, there was no main effect of group but a significant interaction between group and experiment (Table 7.1). Post-hoc independent t-tests showed that checkers were more depressed than non-checkers in the negative-neutral experiment, (Mean (Checker): 6.08±1.20; Mean (Non-checker): 2.25±1.06; \( t(22) = 2.39, p_b = .052, r = .45 \)) but not in the symptom-related experiment (Mean (Checker): 2.75±1.06; Mean (Non-checker): 4.42±1.09; \( t(22) = 1.08, p_b = .57 \)). Interestingly, non-checkers presented a numerically higher score of depression than checkers. There was no significant difference between experiments for checkers and for non-checkers.

In addition, there was no significant difference between groups for the negative-neutral and for the symptom-related experiments, as no significant difference between experiments for checkers and for non-checkers for gender and play for money (Table 7.1). However, the number of videogame players differed between experiments for the non-checker group, where there were more videogame players in the symptom-related (8 yes, 4 no) than in the negative-neutral experiment (3 yes, 9 no) (Table 7.1.). Finally, there was no significant main effect of group, experiment and interaction between group and experiment for age.
**Table 7.1**

*Mean, standard error and statistics for checkers and non-checkers in the negative-neutral and symptom-related experiments, considering all variables analysed for the investigation of sample characteristics. F = female, M = male, Pv = videogame player, NPv = no videogame player, Pm = money player, NPm = no money player; ME = main effect; IA = interaction*

<table>
<thead>
<tr>
<th>Variables</th>
<th>Negative-neutral</th>
<th>Symptom-related</th>
<th>Chi-square or Fisher’s exact test</th>
<th>ME Experiment</th>
<th>ME group</th>
<th>IA group vs. experiment</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Checkers</td>
<td>Non-checker</td>
<td>Checkers</td>
<td>Non-checker</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(between groups)</td>
<td>(between groups)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographic questionnaire (Categorical variables)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>10 F, 2 M</td>
<td>9 F, 3 M</td>
<td>8 F, 4 M</td>
<td>7 F, 5 M</td>
<td>p = 1</td>
<td>p = 1</td>
<td>p = .64</td>
</tr>
<tr>
<td>Play videogame</td>
<td>6 Pv, 6 NPv</td>
<td>3 Pv, 9 NPv</td>
<td>8 Pv, 4 NPv</td>
<td>8 Pv, 4 NPv</td>
<td>p = .4</td>
<td>p = 1</td>
<td>$\chi^2(1, N = 24) = 4.19, p = .04, \Phi = .41$</td>
</tr>
<tr>
<td>Play for money</td>
<td>1 Pm, 11 NPm</td>
<td>0 Pm, 12 NPm</td>
<td>2 Pm, 10 NPm</td>
<td>2 Pm, 10 NPm</td>
<td>p = 1</td>
<td>p = 1</td>
<td>p = .47</td>
</tr>
</tbody>
</table>

| Demographic questionnaire (Continuous variables) |                   |                 | Two-way (2 x 2) ANOVA |               |          |                         |             |
| Variables                           | Negative-neutral | Symptom-related | ME Experiment | ME group | IA group vs. experiment | Effect size |
|                                    | Checkers          | Non-checker     | ME Experiment | ME group | IA group vs. experiment |             |
|                                    | (between groups) | (between groups)|               |          |                         |             |
| Age                                 | 20.25± 0.49       | 22.91± 0.6      | 22.42± 1.14 | 21.25± 0.82 | $F(1, 44) = 0.05, p=.81$ | $F(1, 44) = 3.15, p=.08$ |
### Handedness

<table>
<thead>
<tr>
<th></th>
<th>75.43±11.46</th>
<th>87.27±7.25</th>
<th>80.47±11.56</th>
<th>92.46±3.45</th>
<th>F(1, 44) =0.31 , p=.57</th>
<th>F(1, 44) = 1.72, p=.19</th>
<th>F(1, 44) =0 , p=.99</th>
</tr>
</thead>
</table>

### OCI-R scores

| OCI-R checking | 7.25±0.68  | 0.91±0.25  | 6.75±0.46  | 0.75±0.25  | F(1, 44) = 0.54 , p = .46 | F(1, 44) = 186.49, p < .001** | F(1, 44) = 0.13, p = .71 |
| OCI-R washing | 2.66±0.72  | 1.08±0.68  | 3.75±0.92  | 0.67±0.19  | F(1, 44) =0.24 , p = .62 | F(1, 44) = 11.75 , p < .001** | F(1, 44) = 1.21, p = .27 |
| OCI-R hoarding| 4.50±0.67  | 1.50±0.43  | 5.67± 0.82 | 2.92±0.57  | F(1, 44) =3.73 , p = .06 | F(1, 44) = 18.52, p < .001** | F(1, 44) = 0.03, p = .85 |
| OCI-R ordering| 5.41±0.60  | 2.58±0.63  | 6.5± 0.93  | 3.17±0.81  | F(1, 44) =1.20, p = .27 | F(1, 44) = 16.48, p < .001** | F(1, 44) = 0.10 , p = .74 |
| OCI-R counting| 3.91±0.91  | 0.33±0.14  | 2.98± 0.86 | 1.19±0.34  | F(1, 44) =0.48, p = .48 | F(1, 44) = 23.91, p < .001** | F(1, 44) = 0.32 , p = .57 |
| OCI-R obsessing| 3.91±0.89 | 1.16±0.42  | 4.33±0.79  | 1.33±0.33  | F(1, 44) =0.19 , p = .65 | F(1, 44) = 19.31, p < .001** | F(1, 44) = 0.03, p = .85 |
| OCI-R TOTAL   | 27.66±3.16 | 7.58±1.25  | 31±3.51    | 10±1.90    | F(1, 44) =1.20, p = .27 | F(1, 44) = 61.31, p < .001** | F(1, 44) = 0.03, p = .86 |

### DASS scores

| DASS depression | 6.08±1.29  | 2.25±1.05  | 2.75±1.07  | 4.42±1.10  | F(1, 44) = 0.27 , p = .60 | F(1, 44) = 0.95, p = .33 | F(1, 44) = 6.15, p = .01* |
| DASS anxiety    | 6.08±1.29  | 2.41±0.74  | 5.5±1.07   | 3.5±0.83   | F(1, 44) = 0.6, p = .80 | F(1, 44) = 7.91 , p = .007** | F(1, 44) = 0.68 , p = .41 |
| DASS stress     | 11.16±1.67 | 3.50±1.06  | 8.83±1.58  | 5.42±1.01  | F(1, 44) = 0.02, p = .87 | F(1, 44) = 16.49 , p < .001** | F(1, 44) = 2.42 , p = .12 |

\( \eta^2_{\text{group}} = .80 \)

\( \eta^2_{\text{group}} = .21 \)

\( \eta^2_{\text{group}} = .29 \)

\( \eta^2_{\text{group}} = .27 \)

\( \eta^2_{\text{group}} = .35 \)

\( \eta^2_{\text{group}} = .30 \)

\( \eta^2_{\text{group}} = .58 \)

\( \eta^2_{\text{group}} = .12 \)

\( \eta^2_{\text{group}} = .15 \)

\( \eta^2_{\text{group}} = .27 \)
7.2.3. Laboratory-based Study

7.2.3.1. Laboratory-based study procedure

Upon arrival at the laboratory, participants received the Participant Information Sheet (PIS), containing information about the experiment (Appendix 2.4 for an example of PIS). After reading it, participants should read and sign the Consent form (Appendix 2.5). The total lab-based study took around 2 hours. One booth at the Department of Psychology at University of Surrey was used. It contained the standard Picture Validation task, a binary decision making task, and the standard laboratory-based questionnaires which were the Domain Specific Risk-Taking Scale (DOSPERT); Intolerance of uncertainty questionnaire (IUS); Impulsive Behaviour Scale (UPPS) and the South-Oaks Gambling Screen. Participants also completed a strategy questionnaires (see Chapter 2 for methods details).

The binary decision making tasks were as described in Chapter 2. They only differed in their feedback type. The task was designed to investigate the effect of symptom-related stimuli on RBL (study 6). The specific feedback magnitude for this experiment was negative monetary feedback and symptom-related pictures for the symptom-related experiment. This experiment was compared with the negative-neutral experiment (study 3) (see Chapter 5).

7.2.3.2. Data analysis for the laboratory-based study

Data analysis for study 6 was exactly the same as the one employed at Chapters 5 and 6. For the picture validation task, data analysis considered measures of valence, arousal and comfort. For each measure, a three-way $4 \times 2 \times 2$ mixed ANOVA with the within-participant factor Picture Type (positive, neutral, negative and checking-related pictures) and the between-participant factor Group (checkers vs. non-checkers) and Experiment was conducted for each study. Post-hoc t-tests were conducted and Bonferroni corrections applied when necessary.

Data analysis of the binary decision making task investigated accuracy, sub-block accuracy, probability of shifting after winning and after losing in the previous trial and cross-correlations between the stimulus and the response sequence for the five previous trials (lags). These analysis were described in detail in Chapter 2. The type of statistical test employed for each one of these analysis can be seen at table 7.2. Blocks were analysed
separately because of their maximum accuracy differences and their structural differences. For the exploration / exploitation (win-shift / lose-shift measure), only block 1 was analysed.

Table 7.2.  
*Statistical tests employed for each analysis conducted for the binary decision making task in the symptom-related study*  

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Within-participant factor</th>
<th>Between-participant factor</th>
<th>Type of test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>None</td>
<td>Group</td>
<td>Two-way</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 x 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ANOVA</td>
</tr>
<tr>
<td>Sub-block accuracy</td>
<td>5 Sub-blocks</td>
<td>Checkers vs. non-checkers</td>
<td>Three-way</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Negative-neutral vs. symptom-related</td>
<td>5 x 2 x 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>mixed-ANOVA</td>
</tr>
<tr>
<td>Win-shift</td>
<td>None</td>
<td>negative-neutral vs. symptom-related</td>
<td>Two-way</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Checkers vs. non-checkers</td>
<td>2 x 2</td>
</tr>
<tr>
<td>Lose-shift</td>
<td>None</td>
<td>symptom-related</td>
<td>Two-way</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 x 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ANOVA</td>
</tr>
<tr>
<td>Cross-correlation</td>
<td>5 Lags</td>
<td>5 Lags</td>
<td>Three-way</td>
</tr>
<tr>
<td>each block</td>
<td></td>
<td></td>
<td>5 x 2 x 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>mixed-ANOVA</td>
</tr>
</tbody>
</table>

The lab-based questionnaires were compared between groups and experiments using 2 x 2 ANOVAs with the factors group (checker vs. non-checkers) and Experiment for each study. The exception was the strategy questionnaires that was analysed by using a Pearson Chi-square or Fisher’s exact test to compare groups for each experiment and experiments for each group. No interaction effects were investigated for the strategy questionnaire.
Finally, Kendall-tau correlations were computed for each experiment separately, considering only checkers to investigate if (1) symptom strength recorded with OCI-R total and OCI-R checking subscale scores, (2) comorbidities recorded with the DASS, and (3) measures from the laboratory-based questionnaires are related to behavioural measures recorded in the binary decision making task. These measures were: Accuracies for block 1, block 2, and block 3; win-shift and lose-shift; cross-correlations for lags 1-5 for block 1, 2 and 3.

7.3. Results

7.3.1. Picture Validation task

7.3.1.1. Valence scale

The mixed 4 x 2 x 2 ANOVA with the factors Picture Type, Group and Experiment revealed a significant main effect of picture type on the scores $F(2.64,116.35) = 425.16$, $p < .001$, $\eta^2_p = .90$ but no other effect was significant. The related post-hoc paired t-tests were all statistically significant (all $t(45) > 3.50$, all $p's < .01$). As expected from previous results in Chapters 3 and 5, positive pictures were rated as more positive than neutral, checker-related and negative pictures; neutral pictures were rated as neutral, and more positive than checker-related and negative pictures and checker-related pictures were rated as less negative than negative pictures.

7.3.1.2. Arousal scale

ANOVA results showed only a significant main effect of Picture Type, $F(1.9,83.60)=22.14$, $p < .001$, $\eta^2_p = .33$. Pairwise comparisons were executed to examine this effect. Not all pairwise comparisons were statistically significant and they are shown in table 7.3.
Table 7.3

*Pairwise comparisons for picture type in the picture validation task in the arousal scale for the symptom-related study*

<table>
<thead>
<tr>
<th>Picture-type (Arousal scale)</th>
<th>Mean and standard error</th>
<th>T-test</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive vs. Neutral</td>
<td>4.11±0.21 vs. 5.38±0.13</td>
<td>t(45) = 6.75, ( p_b &lt; .001 )**</td>
<td>r = .70</td>
</tr>
<tr>
<td>Positive vs. Negative</td>
<td>4.11±0.21 vs. 3.91±0.20</td>
<td>t(45) = 0.78, ( p = .44 )</td>
<td></td>
</tr>
<tr>
<td>Positive vs. Checker-related</td>
<td>4.11±0.21 vs. 4.49±0.17</td>
<td>t(45) = 1.62, ( p = .11 )</td>
<td></td>
</tr>
<tr>
<td>Neutral vs. Negative</td>
<td>5.38±0.13 vs. 3.91±0.20</td>
<td>t(45) = 8.14, ( p_b &lt; .001 )**</td>
<td>r = .76</td>
</tr>
<tr>
<td>Neutral vs. Checker-related</td>
<td>5.38±0.13 vs. 4.49±0.17</td>
<td>t(45) = 6.73, ( p_b &lt; .001 )**</td>
<td>r = .70</td>
</tr>
<tr>
<td>Negative vs. Checker-related</td>
<td>3.91±0.20 vs. 4.49±0.17</td>
<td>t(45) = 4.37, ( p_b &lt; .001 )**</td>
<td>r = .54</td>
</tr>
</tbody>
</table>

Similar to previous findings, they showed that positive, negative and checker-related pictures generated more arousal than neutral pictures. Positive and negative pictures and positive and checker-related presented similar amounts of arousal. Checker-related pictures, however, presented less arousal than negative pictures, considering both experiments and groups. No other significant effect was found.

7.3.1.3. **Comfort scale**

Results showed there was a significant main effect of picture type on the scores of comfort \( F(2.81,124.01) = 285.22, p < .001, \eta^2 = .86 \). All post-hoc paired t-tests were statistically significant, as expected (all \( t(45) > 3.50, all \ p's < .01 \)). Basically, positive pictures were rated as more comfortable than neutral, checker-related and negative pictures; neutral pictures were rated as more comfortable than negative and checker-related pictures and checker-related pictures as more comfortable than negative pictures. No other significant effect was found.
In summary, the findings from the picture validation task study reported in Chapters 3 and 5 were only partly confirmed in this study but they were similar to the ones of Chapter 6. In this way, there was no significant difference between checkers and non-checkers in the valence and comfort ratings of the checker-related pictures. This was not expected, as Chapter 3 results showed that checkers usually rate checker-related pictures as more negative than non-checkers but for some reason, this sample rated the checker-related pictures in a similar way. This will be discussed in Chapter 8.

7.3.2. Binary decision making task (study 3 vs. study 6)

7.3.2.1. Accuracy

For each block, a two-way between-participant 2 x 2 ANOVAs with the factors Group (checker vs. non-checker) and Experiment (negative-neutral vs. symptom-related experiment) was conducted to investigate group and experiment differences for each block.

Analysis for block 1 showed there was no significant main effect of group, experiment and interaction experiment vs. group. For block 2, there was no significant main effect of experiment but there was a marginal main effect of group, where checkers presented a higher accuracy than non-checkers ((Mean (Checker): 0.66 ± 0.02; Mean (Non-checker): 0.61± 0.03; $F(1,44) = 3.02, p = .09, \eta^2_p = .06$) and a significant interaction, $F(1,44) = 5.31, p = .02, \eta^2_p = .10$. For block 3, there was a marginal main effect of group, where checkers presented a higher accuracy than non-checkers ((Checker): 0.70 ± 0.02; Mean (Non-checker): 0.65± 0.02; $F(1,44) = 2.88, p = .09, \eta^2_p = .06$) and a marginal main effect of experiment, where participants presented a higher accuracy in the symptom-related experiment than in the negative-neutral experiment, ((Negative-neutral): 0.65 ± 0.02; Mean (Symptom-related): 0.70± 0.02; $F(1,44) = 3.19, p = .08, \eta^2_p = .07$).

Post-hoc independent t-tests showed that checkers (0.68 ± 0.03) had significantly higher accuracy levels than non-checkers (0.55± 0.03; $t(22) = 2.61, p_b=0.03, r=-.48$) in the negative-neutral experiment, what was already shown in Chapter 5. Additionally, no significant group differences were found for the symptom-related experiment (Mean (Checker): 0.63 ± 0.01; Mean (Non-checker): 0.65± 0.03; $t(14.11) = 0.44, p = .66$) (Figure 7.1). Secondly, checkers did not differ between experiments (Mean (Negative-neutral): 0.68 ± 0.03; Mean (Symptom-related): 0.63± 0.01; $t(14.96) = 1.47, p = .16$ and there was also no
significant difference between experiments for non-checkers (Mean (Negative-neutral): 0.56 ± 0.03; Mean (Symptom-related): 0.65± 0.03; \( t(22) = 1.78, p_{b} = .14 \)). (Figure 7.1).

**Figure 7.1** Accuracy mean values for checkers and non-checkers in the negative-neutral and symptom-related experiment for blocks 1, 2 and 3. Figure shows checkers presented a higher accuracy in block 2 than non-checkers for the negative-neutral experiment only.

### 7.3.2.2. Sub-block Accuracy (Learning curve)

Learning curves are displayed in Figure 7.2 for groups and experiments together showing that accuracy improves over time for all blocks. We also investigated group and experimental differences in these learning curves.
For this, a three-way 5 x 2 x 2 mixed-ANOVA was conducted for each block. There was a significant main effect of sub-block for block 1, $F(4.,176) = 15.17, p < .001, \eta^2 = .25$; block 2, $F(4.,176) = 11.69, p < .001, \eta^2 = .21$ and block 3, $F(3.85, 169.85) = 11.46, p < .001, \eta^2 = .20$. Post-hoc paired t-tests are displayed in Table 7.4.
Table 7.4

Pairwise comparisons for the main effect of sub-block accuracy for blocks 1, 2 and 3, considering both experiments and groups together.

<table>
<thead>
<tr>
<th>Sub-block accuracy</th>
<th>Mean and standard error</th>
<th>T-test</th>
<th>Effect size</th>
<th>Mean and standard error</th>
<th>T-test</th>
<th>Effect size</th>
<th>Mean and standard error</th>
<th>T-test</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub-block 1 vs. Sub-block 2</td>
<td>0.50±0.01 vs. 0.56±0.01</td>
<td><em>t</em>(47) = 3.23, pb = .004**</td>
<td>r = .43</td>
<td>0.55±0.02 vs. 0.63±0.02</td>
<td>* &lt; .001**</td>
<td>r = .48</td>
<td>0.59±0.02 vs. 0.68±0.02</td>
<td>* &lt; .001**</td>
<td>r = .59</td>
</tr>
<tr>
<td>Sub-block 1 vs. Sub-block 3</td>
<td>0.50±0.01 vs. 0.59±0.01</td>
<td><em>t</em>(47) = 4.17, pb = .004**</td>
<td>r = .52</td>
<td>0.55±0.02 vs. 0.64±0.02</td>
<td>* &lt; .001**</td>
<td>r = .43</td>
<td>0.59±0.02 vs. 0.70±0.02</td>
<td>* &lt; .001**</td>
<td>r = .60</td>
</tr>
<tr>
<td>Sub-block 1 vs. Sub-block 4</td>
<td>0.50±0.01 vs. 0.63±0.02</td>
<td><em>t</em>(47) = 6.49, pb = .004**</td>
<td>r = .69</td>
<td>0.55±0.02 vs. 0.67±0.02</td>
<td>* &lt; .001**</td>
<td>r = .59</td>
<td>0.59±0.02 vs. 0.69±0.02</td>
<td>* &lt; .001**</td>
<td>r = .56</td>
</tr>
<tr>
<td>Sub-block 1 vs. Sub-block 5</td>
<td>0.50±0.01 vs. 0.64±0.01</td>
<td><em>t</em>(47) = 6.65, pb = .004**</td>
<td>r = .70</td>
<td>0.55±0.02 vs. 0.68±0.02</td>
<td>* &lt; .001**</td>
<td>r = .65</td>
<td>0.59±0.02 vs. 0.70±0.02</td>
<td>* &lt; .001**</td>
<td>r = .52</td>
</tr>
<tr>
<td>Sub-block 2 vs. Sub-block 3</td>
<td>0.56±0.01 vs. 0.59±0.01</td>
<td><em>t</em>(47) = 1.71, pb = .18</td>
<td>r = .81</td>
<td>0.63±0.02 vs. 0.64±0.02</td>
<td>.81</td>
<td>0.68±0.02 vs. 0.70±0.02</td>
<td>.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub-block 2 vs. Sub-block 4</td>
<td>0.56±0.01 vs. 0.63±0.02</td>
<td><em>t</em>(47) = 3.36, pb = .004**</td>
<td>r = .44</td>
<td>0.63±0.02 vs. 0.67±0.02</td>
<td>.08</td>
<td>0.68±0.02 vs. 0.69±0.02</td>
<td>.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub-block 2 vs. Sub-block 5</td>
<td>0.56±0.01 vs. 0.64±0.01</td>
<td><em>t</em>(47) = 4.14, pb = .004**</td>
<td>r = .52</td>
<td>0.63±0.02 vs. 0.68±0.02</td>
<td>.02*</td>
<td>0.68±0.02 vs. 0.70±0.02</td>
<td>.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub-block 3 vs. Sub-block 4</td>
<td>0.59±0.01 vs. 0.63±0.02</td>
<td><em>t</em>(47) = 1.81, pb = .15</td>
<td>r = .35</td>
<td>0.64±0.02 vs. 0.67±0.02</td>
<td>.13</td>
<td>0.64±0.02 vs. 0.69±0.02</td>
<td>.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub-block 3 vs. Sub-block 5</td>
<td>0.59±0.01 vs. 0.64±0.01</td>
<td><em>t</em>(47) = 2.45, pb = .03*</td>
<td>r = .34</td>
<td>0.64±0.02 vs. 0.68±0.02</td>
<td>.06</td>
<td>0.64±0.02 vs. 0.69±0.02</td>
<td>.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub-block 4 vs. Sub-block 5</td>
<td>0.63±0.02 vs. 0.64±0.01</td>
<td><em>t</em>(47) = 0.85, pb = .60</td>
<td>r = .96</td>
<td>0.67±0.02 vs. 0.68±0.02</td>
<td>.69</td>
<td>0.69±0.02 vs. 0.69±0.02</td>
<td>.96</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
These post-hoc tests revealed that participants were able to increase performance in all the three blocks; as accuracy levels progressively increased across sub-blocks. (Figure 7.2). Learning in block 1 was different from blocks 2 and 3, as previously shown in Chapters 5 and 6. In block 1, participants needed at least two sub-blocks to increase accuracy and accuracy progressively increased until sub-block 4, while in blocks 2 and 3, participants only needed the first sub-block (from sub-block 1 to 2) to learn the task (Figure 7.2). However, while accuracy values were kept constant in block 3, participants were able to continue to increase accuracy in block 2 until sub-block 5 (sub-block 2 vs. sub-block 5 in table 7.4).

There was a significant interaction between group and sub-block, $F(4.,176) = 2.97, p = .02, \eta^2_p = .06$ and a significant interaction between experiment and sub-block, $F(4.,176) = 2.62, p = .03, \eta^2_p = .05$ for block 2. To furtherly investigate these values, five independent t-tests were conducted between groups for all experiments and five independent t-tests were executed between experiments for all participants together.

Analysis revealed that there was a significant difference between groups in sub-block 5 of block 2, where checkers presented a higher accuracy than non-checkers in this sub-block ((Mean (Checker): 0.74 ± 0.03; Mean (Non-checker): 0.62± 0.03; $t(44.73) = 2.77, p_b = 0.04, r= .38$) (Figure 7.3). There was no difference between experiments for all the sub-blocks, when considering Bonferroni corrections. Finally, there was no significant interaction between experiment and sub-block accuracy for blocks 1, and 3, no significant interaction between group and sub-block accuracy for blocks 1 and 3 and no significant interaction between experiment, group and sub-block for all the blocks, including block 2.
In summary, the accuracy analysis for each block separately revealed, firstly, that learning occurred in all three blocks. Note that it was slower in block 1, however. Secondly, there were no group or experimental differences for the random sequence block (block 1) and for block 3, where target option maximisation and pattern search strategies were combined. Finally, no experimental differences were found for block 2 but there was a group difference for this block, where checkers presented a higher accuracy in sub-block 5 than non-checkers. These findings will be discussed later on in Chapter 7.4.

7.3.2.3. Exploration / exploitation: Win-stay / lose-shift measures in Block 1

Figure 7.4 indicates that checkers performed similarly in the negative-neutral experiment and in the symptom-related experiment, as well as non-checkers. Indeed, this figure again indicates that checkers and non-checkers are not doing excessive exploration (more win-shift and more lose-shift) because only two participants presented a higher probability of shifting after losing and winning in the negative-neutral experiment (see right superior quadrant for both experiments in figure 7.4). Besides, confirming previous findings, it is possible to see that both groups prefer to shift after losing, as only one participant is seen at the right inferior quadrant of figure 7.4 (more win-shift/lose-stay) in the negative-neutral experiment, where the probability of shifting after losing is the lowest. Additionally, checkers and non-checkers seem to be similarly shifting after losing and winning, as the majority of
participants are concentrated between the left superior and left inferior quadrant (see left superior and inferior quadrant of figure 7.4).

\[ \text{Figure 7.4 Win-shift / Lose-shift strategies per participant divided by group and experiment.} \]

The left and inferior quadrant of the figure indicates more exploitation after winning and after losing (win-stay / lose-stay). The right and superior quadrant indicate more exploration after winning and losing (win-shift / lose-shift). The left and superior quadrant indicates more exploration after losing only (win-stay / lose-shift). The right and inferior quadrant indicates more exploration after winning only (win-shift / lose-stay). Figure indicates that checkers and non-checkers similarly performed in the negative-neutral and symptom-related experiment.

Results from the two-way 2 x 2 ANOVA confirmed this interpretation, showing there was no significant main effect of experiment, group and interaction between experiment and group for win-shift and for lose-shift. In this way, the probability of shifting after winning and losing is similar for both checkers and non-checkers. Additionally, the probability of shifting after winning and losing for checkers and non-checkers is similar between the negative-neutral and symptom-related experiment.
7.3.2.4. Cross-correlation

7.3.2.4.1. Cross-correlations for Block 1

Figure 7.5 shows the auto- and cross-correlations for block 1. Here, the difference between auto-and cross-correlation scores seems to be most evident for lag 1, but maybe also present for lags 2 and 3. The three-way 5 x 2 x 2 mixed ANOVA on the auto- and cross-correlation difference scores revealed a significant main effect of lag, $F(4, 176) = 12.22, p < .001, \eta^2 = .21$. This was caused by a significantly higher difference between the participant’s auto- and cross-correlation values for lag 1 than for lags 2, 3, 4, and 5, all $t(45) > 3.50$, all $p$’s < .01. This might be because participants assumed a greater dependency between the current and the previous event (lag 1) than present in a random sequence, as it has happened in previous chapters. In addition, the difference between auto-and cross-correlations was similar but with opposite values at lag 2 compared to lag 3, ((Mean (lag 2): -0.04 ± 0.02; Mean (lag 3): 0.04± 0.02; $t(47) = 4.60$, $p_b < .001$, $r=0.56$), and was higher at lag 3 compared to lag 4 ((Mean (lag 3): 0.04 ± 0.02; Mean (lag 4): -0.02± 0.02; $t(47) = 3.02$, $p_b = .04$, $r=0.40$). This effect might be more due to uncontrolled changes in the experimental sequences, because the experimental sequence score is slightly enhanced at lag 3 and because this difference was not found for the others pairwise comparisons of lag 2 (2-4, 2-5) or lag 3 (3-4, 3-5). In this way, supporting findings from Chapters 5 and 6, this data supports that participants employ the immediate previous results to make predictions (Figure 7.5).

![Figure 7.5](image-url)
together for the symptom-related study. Figure indicates that participants correlated more than expected in lag 1, when considering the experimental sequence correlation values..

7.3.2.4.2. Cross-correlations for Block 2

Figure 7.6 shows that both groups generally followed the pattern in the experimental sequence. In this way, checkers and non-checkers were able to learn sequential effects in this block. As expected the absolute cross-correlation values were not as large as the auto-correlation values. Figure 7.6 also reveals group differences in relation to the participant responses and the pattern sequence, where checkers performing in the negative-neutral experiment seem to be closer to the auto-correlation values than non-checkers.

![Figure 7.6 Cross-correlation probability from lags 1 to 5 between checkers and non-checkers for the negative-neutral and symptom-related experiment in block 2. The averaged correlation values of the experimental sequence are shown in black. Figure shows that checkers cross-correlated more than non-checkers in lags 2 and 4 for the negative-neutral experiment.](image)

Actually, the three-way 5 x 2 x 2 ANOVA for block 2 showed there was a significant interaction of lags vs. experiment vs. group, $F(1.71,75.46) = 4.93, p = .01, \eta^2_p = .10$ and a main effect of lag, $F(1.71,75.46) = 57.37, p < .001, \eta^2_p = .56$, that was expected as lag values
were different in the sequence, due to the presence of sequential effects. There was no significant effects of lags vs. experiment and lags vs. group.

For the significant interaction, five two-way 2 x 2 ANOVAs from lags 1 to 5, considering groups and experiments as between-participant factors were performed. Results showed a significant interaction between experiment and group for lag 2, $F(1, 44) = 6.77$, $p = .01$, $\eta^2 = .13$, lag 4, $F(1, 44) = 5.81$, $p = .02$, $\eta^2 = .11$ and a marginal interaction for lag 5, $F(1, 44) = 3.13$, $p = .08$, $\eta^2 = .06$. Independent t-tests were conducted for lags 2, 4 and 5 separately, considering the investigation of differences between groups for each experiment and differences between experiments for each group (table 7.5).

Table 7.5

**Independent t-tests for lags 2, 4 and 5 for block 2 between groups and between experiments**

| Lags | Between groups analysis |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Negative-neutral |  |  | Symptom-related |  |  |  |
| Checkers | Non-checkers | Checkers | Non-checkers |  |  |  |  |
| Lag 2 | -0.20±0.08 | -0.57±0.09 | -0.38±0.04 | -0.31±0.01 | $t(22) = 2.89$, $p_b = .02^*$, $r = 0.52$ |  |  |
| Lag 4 | 0.13±0.06 | 0.39±0.07 | 0.25±0.04 | 0.19±0.08 | $t(22) = 2.60$, $p_b = .03^*$, $r = 0.48$ |  |  |
| Lag 5 | -0.14±0.03 | 0.18±0.03 | 0.19±0.02 | 0.12±0.03 |  | $t(22) = 0.84$, $p = .40$ |  |
| Between experiment analysis |  |  |  |  |  |  |  |
| Checkers | Negative-neutral | Symptom-related |  | Non-checkers | Negative-neutral | Symptom-related |
| Lag 2 | -0.20±0.08 | -0.38±0.04 | -0.57±0.09 | -0.31±0.01 | $t(22) = 1.85$, $p_b = .16$ | $t(22) = 1.90$, $p_b = .14$ |
| Lag 4 | 0.13±0.06 | 0.25±0.04 | 0.42±0.06 | 0.19±0.08 | $t(22) = 1.46$, $p_b = .15$ |  |
| Lag 5 | -0.14±0.03 | 0.19±0.02 | 0.18±0.03 | 0.12±0.03 |  | $t(22) = 1.10$, $p = .27$ |  |
|  |  |  |  |  |  |  |  |

232
Basically, checkers were closer than non-checkers to the experimental sequence in lags 2 and 4 in the negative-neutral experiment. This effect was already reported in Chapter 5. There was no significant effect of group for lag 5 and no effect of group for any of the lags for the symptom-related experiment. Additionally, there was no effect of experiment for checkers and for non-checkers (Figure 7.6).

7.3.2.4.3. Cross-correlations for Block 3

Figure 7.7 shows cross-correlation values for block 3, considering both experiments together. The three-way 5 x 2 x 2 ANOVA for block 3 showed there was a significant effect of lags vs. group, $F(1.50.,66.10) = 3.88$, $p = .03$, $\eta^2 = .08$ and there was a main effect of lag, $F(1.50.,66.10) = 46.76$, $p < .001$, $\eta^2 = .51$, that was already expected as lag values were different in the sequence, due to the presence of sequential effects.

The significant effect of lags vs. group was investigated with five independent t-tests for each lag. However, analysis revealed that there was no significant main effect of group for any of the lags analysed.

*Figure 7.7* Cross-correlation probability from lags 1 to 5 between checkers and non-checkers considering both experiments together in block 3. The averaged correlation values of the experimental sequence are shown in black. Figure indicates that checkers and non-checkers presented similar cross-correlation values for all lags.
In summary, the findings from the auto- and cross-correlation analysis revealed that no group or experimental differences were found for block 1 (random block with target option probability manipulation) and block 3 (target option and conditional probability manipulation). Group differences were revealed for block 2, as checkers learned the pattern in block 2 better in the negative-neutral experiment compared to non-checkers. This findings were already expected and revealed at Chapter 5. However, groups were similar in the symptom-related experiment and there was no experiment difference for checkers and non-checkers. These findings will be discussed below.

7.3.3. Laboratory-based questionnaires

7.3.3.1. Intolerance of uncertainty scale (IUS)

The two-way 2 x 2 ANOVA analysis showed that there was a significant main effect of group for factor 1 ((Mean (checkers): 37.75 ± 2.02; Mean (non-checkers): 30.79± 2.37; \(F(1,44) = 5.00, p = .03, \eta^2 = .10\)); factor 2 ((Mean (checkers): 37.46 ± 1.99; Mean (non-checkers): 26.71± 2.03; \(F(1,44) = 13.98, p = .001, \eta^2 = .24\)) and total score ((Mean (checkers): 75.21 ± 3.59; Mean (non-checkers): 57.5± 4.30; \(F(1,44) = 9.88, p = .003, \eta^2 = .18\)), where checkers presented a higher score of intolerance of uncertainty than non-checkers. There was no significant main effect of experiment and no interaction between experiment and group for factor 1 (Uncertainty has negative behavioural and self-referent implications), factor 2 (uncertainty is unfair and spoils everything) and total score.

7.3.3.2. Domain Specific Risk-Taking Scale and Risk-Perception scale (DOSPERT)

The two-way 2 x 2 ANOVA analysis revealed there was a significant main effect of experiment for risk-taking finance, ((Mean (negative-neutral): 13.66 ± 1.31; Mean (symptom-related): 19.29± 1.54; \(F(1,44) = 8.37, p = .006, \eta^2 = .16\)), where participants risked less in the negative-neutral than in the symptom-related experiment. Additionally, there was a significant main effect of experiment for risk-perception finance ((Mean (negative-neutral): 33.38 ± 1.21; Mean (symptom-related): 29.13± 1.42; \(F(1,44) = 5.36, p = .02, \eta^2 = .11\)) and risk-perception total score ((Mean (negative-neutral): 140.17 ± 3.24; Mean (symptom-related): 128.71± 3.65; \(F(1,44) = 5.27, p = .02, \eta^2 = .10\)), where participants presented a higher perception of risk in the negative-neutral experiment in comparison with the symptom-related experiment. There
was no significant main effect of group or interaction for risk-taking finance, risk-taking total, risk-perception finance and risk-perception total.

7.3.3.3. **UPPS Impulsivity Scale**

There was no significant main effect of experiment, group and interaction for all the UPPS sub-sections and total score. Thus, this indicates that checkers and non-checkers presented similar values of impulsivity.

7.3.3.4. **South Oaks Gambling Screen**

There was no significant main effect of group, experiment and interaction for gambling in the South-Oaks screen.

7.3.3.5. **Strategy questionnaires**

Categorical data from the strategy questionnaire showed that checkers employ strategies and visualize patterns in the same way between experiments in blocks 1, 2 and 3. In the other hand, non-checkers employed strategies and visualized patterns in the same way in blocks 1 and 2 but there was a significant difference in the visualization of patterns in block 3, where non-checkers saw more patterns in the symptom-related experiment (11 yes, 0 no) than in the negative-neutral experiment (5 yes, 7 no), $p = .005$.

Regarding group differences, there was no significant group difference for the negative-neutral and for the symptom-related experiment for blocks 1, 2 and 3 for the use of a strategy. For the visualization of a pattern, there was a significant difference between groups in the symptom-related experiment, where checkers saw less patterns (3 yes, 9 no) than non-checkers (9 yes, 2 no) in block 1. Additionally, there was a marginal effect of group for the visualization of a pattern for block 2, where checkers saw more patterns (11 yes, 1 no) than non-checkers (6 yes, 6 no), $p = .07$ in the negative-neutral experiment. This effect was already reported in Chapter 5.

In summary, the analysis of the lab-based questionnaires revealed that checkers showed greater levels of intolerance of uncertainty. In addition, non-checkers saw more patterns in block 1 than checkers in the symptom-related experiment and they saw more patterns in the symptom-related experiment than in the negative-neutral experiment in block
3. Finally, checkers saw more patterns than non-checkers in the negative-neutral experiment in block 2.

7.3.4. Correlation analysis between behavioural measures from the binary decision making task and laboratory-based questionnaire scores

7.3.4.1. Negative-neutral experiment (Study 3)

Results were already reported in Chapter 5 and 6 for the negative-neutral experiment and revealed that symptom severity (checking OCI-R) was correlated with performance as reflected in cross-correlation values in block 2 (Dif lag 4: $r_\tau = .46$, $p = .04$), indicating that symptom severity decrease pattern learning in block 2. In addition, comorbidity levels did not correlate with the binary decision making task performance. Finally, laboratory-based questionnaire ratings correlate with accuracy in block 2 in the binary decision making task (Factor 1: $r_\tau = -.61$, $p = .006$; Factor 2: $r_\tau = -.47$, $p = .03$; Total score: $r_\tau = -.54$, $p = .016$) and pattern learning in lag 2 (Factor 1: $r_\tau = -.54$, $p = .01$; Factor 2: $r_\tau = -.46$, $p = .03$; Total score: $r_\tau = -.46$, $p = .03$). Indeed, this analysis indicates that higher levels of intolerance of uncertainty were related to a decreased performance in the binary decision making task for block 2. In this experiment, intolerance of uncertainty is influencing behaviour.

7.3.4.2. Symptom-related experiment (Study 6)

Kendall tau correlation analysis for the symptom-related experiment revealed that there was a significant and negative correlation between checker OCI-R and accuracy in block 2, $r_\tau = .47$, $p = .05$, where checkers with higher symptoms of checking seem to present a lower accuracy in block 2. In this way, more checking symptoms decrease performance in block 2.

Additionally, there was a significant and positive correlation between intolerance of uncertainty for factor 1 and accuracy in block 1, $r_\tau = .52$, $p = .02$, where checkers with higher intolerance of uncertainty presented a higher accuracy in block 1. In this way, checkers with higher intolerance of uncertainty presented a higher performance in block 1.

There was also two significant correlations between lag 5 of block 2 and DASS depression, $r_\tau = .53$, $p = .02^*$ and anxiety $r_\tau = .57$, $p = .01$, where checkers who were closer to the experimental sequence in lag 5 of block 2, present lower values of depression and anxiety in checkers. In this way, it seems depression and anxiety are influencing performance in block 2.
2, as more depression and anxiety could increase the Difference between the experimental sequence and the participant’s sequence in lag 5 for block 2. New studies should be developed to investigate the relationship of depression and anxiety and performance in RBLTs with patterns in checkers.

Summing up, more checking symptoms decrease performance in block 2 in the symptom-related task. This is also the case for the negative-neutral experiment, where higher checking symptoms decrease performance in the pattern of block 2. In addition, higher intolerance of uncertainty increases performance in block 1, what is the opposite for the negative-neutral experiment, where higher intolerance of uncertainty decreases performance in checkers but for block 2 only. Finally, more depression and anxiety in checkers increase the Difference between the experimental sequence and the participant’s sequence in lag 5 for block 2, so more depression and anxiety is decreasing pattern learning in block 2. There was no significant correlation between the comorbidity measures depression and anxiety and the binary decision making task measures in the negative-neutral experiment.

7.4. Discussion

Previous studies revealed that individuals with OCD have emotional regulation and cognitive deficits in decision making (Starcke et al., 2010). More specifically, OCD participants present deficits in decision making tasks under ambiguity which are often equivalent to probabilistic reward-based learning tasks (Starcke et al., 2010; Kim et al., 2015; Zhang et al., 2015). As these tasks usually examine distinct factors at the same time, as feedback direction and feedback magnitude, the present thesis designed a binary decision making task where feedback direction could be separately investigated from feedback magnitude. Indeed, the study from Chapter 5 examined the effect of the manipulation of feedback direction in the RBL of subclinical checkers and studies from Chapter 6 investigated the effect of the manipulation of feedback magnitude in the RBL of subclinical checkers within a positive feedback direction (Positive study) and within a negative feedback direction (Negative study). Actually, these studies showed that subclinical checkers seem to present a bias towards exploitation when exposed to specific values of error magnitudes. Data also indicated that checkers were unable to adapt after being exposed to distinct error magnitude values, performing the same bias towards exploitation in these distinct tasks. However, it was not clear if specific error magnitudes associated with symptom-related stimuli would produce the same bias or would increase it in subclinical checkers.
In this way, the presented study investigated subclinical checkers and matched non-checkers and compared their performance in two separate experiments that manipulated the presence or the absence of symptom-related aversive stimuli in binary decision making tasks disguised as a binary decision making task. One experiment used negative monetary plus neutral emotional pictures vs. absent penalties (negative-neutral experiment as study 3) and the second experiment used negative monetary plus symptom-related emotional pictures vs absent penalties (symptom-related experiment as study 6). In both experiments, three blocks were present, i.e. a random sequence block with a target-option manipulation, a pattern sequence block with equal winning probabilities for both target options, and a pattern sequence block where one target option had a higher chance of winning.

The first aim of this study was to examine the effect of the presence of symptom-related emotional stimuli, (i.e. general aversive stimuli vs. symptom-related stimuli) in people with subclinical checking symptoms and compare them to non-checkers in a random sequence block and to compare our findings with the previously reported studies. It was expected that (1) symptom-related stimuli would distinctively affect subclinical checkers and, due to its aversive nature for subclinical checkers, this would increase the bias towards exploitation in comparison with the negative-neutral experiment OR that (2) symptom-related stimuli would not increase the bias towards exploitation in subclinical checkers. This would be the case because the influence of symptom-related stimuli would only depend on the feedback magnitude associated with the stimuli but not with its specificity as symptom-related stimuli. In other words, a symptom-related stimuli could or could not affect subclinical checker decision making, providing that the error magnitude associated with it presents a specific magnitude that affects subclinical checkers.

Based on previous findings from the Negative study at Chapter 6, increasing the feedback magnitude between the negative-neutral and negative-negative experiment did not affect the bias towards exploitation in checkers, so increasing the feedback magnitude between the negative-neutral and the symptom-related experiment, would also do not affect the bias towards exploitation in subclinical checkers, as it was hypothesized that subclinical checkers would be unable to adapt the bias towards exploration/exploitation when exposed to values of error magnitudes that surpass a specific threshold.

The findings from the accuracy analysis of all three blocks showed that participants were able to learn the task over time, as it was previously shown in Chapters 5 and 6. For block 1, accuracy levels increased from sub-block 1 to sub-block 3 and stabilised thereafter.
In other words, participants needed two sub-blocks (64 trials) to increase their performance. For blocks 2 and 3, participants only needed one sub-block for learning. Actually, this could be because block 1 was more difficult than blocks 2 and 3 (lower maximum accuracy), and it did not present a pattern. Finally, results also indicated that checkers were able to learn the task, regardless of the presence of symptom-related stimuli, so this seems to do not affect checkers differently. In addition, there was no significant difference in learning between the two experiments for checkers, so symptom-related stimuli might not differently affect learning in checkers in comparison with a general aversive stimuli.

More importantly, no group or experimental accuracy differences were found for the random sequence block 1 with differential target option probabilities. The findings from the auto- and cross-correlation analysis support these results. As expected from our previous studies, only a main effect of lags was found for block 1, indicating that participants were being influenced by the immediately preceding trial to make decisions, a concept that supports first order effects in RBL (Wilder et al., 2009). Generally, this data confirms that subclinical checkers are as equally able to learn the random sequence task as non-checkers in the negative-neutral and the symptom-related experiment. This corroborates the hypothesis that the presence of symptom-related stimuli do not affect checker performance differently from general aversive stimuli.

In addition to the accuracy analysis in the random sequence block, the usage of exploration /exploitation strategies was evaluated but there was no exploration / exploitation differences between the subclinical checker group and the non-checker group in both experiments, and also no exploration / exploitation differences between experiments. This indicates that checkers and non-checkers are similarly influenced by errors and winnings in both tasks and this could indicate (1) the presence of symptom-related stimuli does not distinctively affect checkers from general aversive stimuli and (2) checkers are unable to adapt the bias between exploration / exploitation when the value of the error magnitude of one experiment is higher than the one at the negative-neutral experiment. In fact, the error magnitude of the symptom-related experiment is an intermediate value between the negative-neutral and the negative-negative experiment and it is higher than the one of the negative-neutral condition, so increasing the error magnitude of the task did not affect checkers, as in the Negative study (negative-neutral vs. negative-negative).

Interestingly, results are different for non-checkers. Actually, non-checkers saw more patterns in block 1 than checkers in the symptom-related experiment, what could indicate that
non-checkers are more biased towards exploration in this experiment than checkers, corroborating results from the Negative study, where non-checkers presented more exploration than checkers in the experiment with higher error magnitude (negative-negative). Note that block 1 did not present a pattern, so seeing more patterns here could indicate non-checkers are doing more exploration, possibly searching for inexistent patterns in this sequence.

Interestingly, there were no group and experimental differences for non-checkers in this study as there was in the Negative study in Chapter 6 (negative-neutral vs. negative-negative experiment) for lose-shift. If non-checkers are being influenced by the error magnitude of the symptom-related experiment and enhancing their bias towards exploration as in the negative-negative experiment, an experimental difference for lose-shift would be expected for this group. This was not the case. However, it is possible that there were no findings here because the difference of the error magnitude between the negative-neutral and symptom-related experiment was not enough to produce an effect in lose-shift (see the Negative study, Chapter 6). Indeed, the error magnitude difference between the negative-neutral and the negative-negative was higher than the difference between the negative-neutral and the symptom-related experiments. Clearly, if this is correct, a new study comparing the negative-neutral experiment with a negative experiment with an even higher error magnitude would present similar effects as the Negative study, as the difference between error magnitude values would be higher than the ones of the negative-negative study. Additionally, a new study comparing the negative-neutral with a negative experiment with a similar error magnitude value of the symptom-related experiment might present no effects for non-checkers;

Regarding correlations between the performance in the random sequence block and the questionnaires, results revealed that (1) financial risk-taking is related to more shifts after winning in the negative neutral experiment, which is related to higher auto- and cross-correlation differences in lag 1 because risk-takers seem to be more influenced by previous winnings. (2) Checkers with higher intolerance of uncertainty presented a higher performance in block 1 in the symptom-related experiment. These correlations were not common between the negative-neutral and symptom-related experiment, so they could not be used to explain the results. However, there was a significant correlation between intolerance of uncertainty and block 1 accuracy. This was only found for the symptom-related experiment but distinct correlations were also found between intolerance of uncertainty and the binary decision
making variables in other experiments within the negative feedback direction, so this could indicate that intolerance of uncertainty is influenced or influences subclinical checker performance within a negative feedback direction.

Regarding results from block 2, there was a significant interaction of experiment vs. group but this difference only regarded the negative-neutral experiment, where checkers presented a higher accuracy than non-checkers (see Chapter 5 for more details). In fact, this difference could be also reflecting the significant difference between checkers and non-checkers in sub-block 5 for block 2. Actually, checkers also presented a higher accuracy for this sub-block than non-checkers but only if considering both experiments together. In this way, it is possible that this difference is only reflecting the differences in accuracy found for the negative-neutral experiment in Chapter 5 and this is actually the case for the cross-correlation findings for block 2, where there was a significant interaction between groups and experiments but there was no group difference for the symptom-related experiment and no experimental difference for the checker and non-checker group. However, again, there was a significant effect of group for the negative-neutral experiment that was already revealed at Chapter 5.

For block 3, there was no significant effect of group and experiment for accuracy, win-shift and lose-shift and auto- and cross-correlation measurements and this actually corroborates previous studies in Chapters 5 and 6, where no effects were found for this block. Nevertheless, while it seems there was no effect of group or experiments for the symptom-related task, the strategy questionnaire analysis showed that non-checkers were able to see more patterns in block 3 in the symptom-related than in the negative-neutral experiment. This is interesting because it suggests that the symptom-related experiment optimal strategy is easier to find than the negative-neutral experiment but only for non-checkers. This could indicate that non-checkers are more biased towards exploitation, ignoring disruptions in this pattern or, in the other hand, this could indicate that non-checkers, due to more exploration, find the pattern easily. In fact, the strategy questionnaire data alone, with no other significant finding regarding the binary decision making task, is not enough to support an interpretation of more exploration or more exploitation in the task by non-checkers. However, these results support findings from the Negative study in Chapter 6, where non-checkers seemed to see more patterns in block 3 for the experiment with higher error magnitude (negative-negative experiment) than for the experiment with lower error magnitude (negative-neutral experiment). In this way, this supports the idea that non-checkers see more patterns in the
sequence when the error magnitude of the task increases. Nonetheless, it is not possible to know if this is because they increase their bias towards exploitation in sequences with patterns. If this is the case, this would indicate that non-checkers are more biased towards exploration in sequences without patterns (block 1) and more biased towards exploitation in sequences with patterns (block 3), when the error magnitude of the task increases. Note, however, that checkers, might not present this ability to adapt the bias towards exploration and exploitation when the error magnitude of the task increases. In fact, even for the sequences with higher maximum accuracies (blocks 2 and 3) and consequently easier to learn, the bias towards exploitation persists in checkers, as no differences of accuracy and cross-correlation were found for block 2 between the negative-neutral and the symptom-related experiment or between the negative-neutral and the negative-negative experiment.

Finally, in relation to the Kendall-tau correlations, there were no common correlations between the two experiments that could explain results. Actually, the only common variable was intolerance of uncertainty. Again, this could indicate that intolerance of uncertainty plays an important role when checkers are exposed to experiments within a negative direction feedback or that values of intolerance of uncertainty are influenced by some common factor, regarding these tasks.

In summary, results have shown that checkers are not differently affected by symptom-related stimuli in comparison to general aversive stimuli, while non-checkers might be. This indicates that the error magnitude of the emotional stimuli is influencing RBL in checkers and non-checkers and not its nature as symptom-related stimuli. This supports the hypothesis that checkers are not able to adapt their decision making behaviour when exposed to higher values of error magnitudes that surpass a specific value (maybe below or equal to the negative-neutral experiment). Note that it is not because the negative stimuli is related with symptom-related stimuli that this could differently affect checker’s behaviour. In fact, it is possible that symptom-related stimuli might affect checkers, depending on its error magnitude value but not because they are related with checking symptoms. Actually, it seems that there is a threshold of error magnitude that bias checkers towards exploitation. After this threshold is surpassed, checkers might not be able to use exploration. Clearly, it is possible that this distinct behaviour between checkers and non-checkers is related with negative prediction error deficits in OCD, so they would be unable to correct hypothesis after an error and consequently they would not be able to explore for new strategy hypothesis. This will be furtherly discussed in Chapter 8.
7.4.1. Limitations and future directions

It is important to consider that the sample employed in this study presented some particularities. Actually, different from our previous studies, there was no group difference in the symptom-related experiment, regarding DASS depression, where checkers usually presented more depression than non-checkers in the DASS questionnaire. Note that non-checkers presented a numerically but not significant DASS depression score than checkers. Additionally, there was an effect of videogame experience for non-checkers, as there were more non-checker videogame players in the symptom-related experiment than in the negative-neutral experiment. Moreover, checkers and non-checkers presented the same amount of gambling habits in the South-Oaks questionnaire, what was different from our previous studies, where checkers usually gambled more than non-checkers. Finally, there was no difference between groups for the valence and comfort of the checking-related pictures. Actually, this does not contradict results from Chapter 6 but for this specific study, it was aimed that checkers would rate symptom-related pictures as more negative than non-checkers, as it was shown in Chapter 3. Indeed, checkers were more negatively affected by these pictures here, as their general valence and comfort ratings were more negative than the neutral pictures. The problem is that non-checkers rated these pictures as too negative and not neutral, as it was previously aimed, decreasing the difference between groups.

In truth, this is possibly related to the fact that the sample employed in this study might be different from the others samples from Chapters 3, 5 and 6, as shown by results of the DASS depression, videogame experience and gambling behaviour in the South-Oaks. In this way, it is important that future studies could replicate this study with a new sample, aiming either for a significant group difference in the rating of the symptom-related pictures and for a lower rating of depression in the non-checker group. However, while this is not done, it is still possible to discuss the results found for this study, considering the symptom-related experiment as an experiment with an intermediate value of error magnitude between the negative-neutral experiment and the negative-negative experiment. Additionally, it is still possible to investigate how checkers deal with these checker-related pictures in the symptom-related experiment in comparison with a general negative experiment, as the negative-neutral experiment.
7.5. **Conclusion**

Symptom-related stimuli seems to not influence subclinical checkers in a particular way. Indeed, it is possible that the error magnitude of these stimuli is more important than the nature of the stimuli. Based on that, while non-checkers are able to adapt their bias towards exploration / exploitation in a RBLT within distinct magnitudes of error, checkers might be unable to adapt when the error magnitude of the emotional stimuli in the task is above a certain threshold.
Chapter 8 – General Discussion

The aim of this thesis was to investigate decision making deficits in obsessive-compulsive disorder (OCD). In particular, this thesis has focused on the investigation of implicit reward-based learning deficits in subclinical checkers as previous research has identified specific deficits in OCD, regarding implicit RBLTs but not others decision making tasks where the target probabilities are explicit and previously known by participants (decision making under risk). Usually, studies trying to investigate reward-based learning deficits in OCD employ decision making tasks that investigate several reward related factors at the same time, as feedback direction (positive vs. negative), feedback magnitude (more positive vs. less positive) or feedback delay (feedback presentation immediately after the choice vs. delayed after the choice). This makes difficult to interpret about the specific deficits regarding RBLTs in OCD. In this way, this investigation helped to disentangle these different factors in RBL in a systematic way, where feedback direction and feedback magnitude were investigated throughout the use of a binary RBL decision making task.

Additionally, this thesis focused on the investigation of the differential effects of feedback magnitude alterations, when considering symptom-related stimuli (checking pictures) (study 6) and on the effect of systematic changes in target probabilities vs. conditional probabilities on learning in the RBLTs (random event sequence vs. pattern learning) as previous studies usually ignored the effect of the manipulation of conditional probabilities in RBLTs in OCD. Further, the results of this thesis will be discussed in terms of their contribution, limitations and how they are relevant in future experimental design. The discussion will close with the presentation of final conclusions.

8.1. Synthesis of research findings

The findings for the online questionnaire confirmed that our sample was characterized by checkers and non-checkers, both with mild depression and moderate anxiety and depression. Checkers usually presented higher levels of depression, anxiety and stress but sometimes this difference was inexistent, as in the case of depression in the symptom-related study. Generally, there was no group difference regarding gender, videogame experience and gambling experience (play for money). Age differences between groups were sometimes present but the difference was not enough to affect decision making strategies, as groups were always inside the 18-22 years old age group (Victorino et al., unpublished). Importantly,
this investigation was successful to exclude common comorbidities from the sample, as clinical depression, anxiety, panic disorder and PTSD.

Others findings in this thesis confirmed that participants were unmedicated, as they were part of a subclinical sample and their symptoms were mainly characterised by checking, even though others symptoms regarding others OCD subtypes were generally present in the checker group.

The findings regarding the Picture validation task study are very similar to the ones found in Simon et al. (2010) study, where obsessive compulsive individuals usually rated symptom-related and general aversive emotional stimuli as more negative than the control group. These results were both true for the valence scale and for the new comfort scale. Unfortunately, these results were not replicated in posterior samples employed in the subsequent binary decision making studies. Specially, checker-related pictures were usually rated similarly between the checker and the non-checker group, even in the symptom-related study, where it was necessary to find distinct ratings between groups for the checker-related pictures. Actually, a difference between groups for the checker-related pictures were only found again in the neutral study (Chapter 5), where checkers rated these pictures as more uncomfortable than non-checkers. There was no difference for valence and for the general aversive pictures, though. In this way, it seems this difference is very dependent on the sample employed in the study in question and it also depends on the sample size, as when pooling all samples together, group differences for checking-related valence and comfort were found again. As studies from Chapters 5 and 6 did not employ the checker-related pictures in the binary decision making task, this lack of difference did not constitute a problem but for the symptom-related experiment, it would be important that this study would be replicated with a new sample in the future, especially because this sample also presented some odd characteristics, as more videogame experience in the non-checker group in the symptom-related experiment, no group differences in gambling habits for the South-Oaks questionnaire and numerically higher levels of depression in the DASS for non-checkers. However, even though this study should be replicated, results have shown that these pictures were rated as more negative than the neutral pictures and less negative than the negative pictures in both groups, so the symptom-related study still presents an intermediate value of error magnitude between the negative-neutral and the negative-negative experiment. In this way, the symptom-related experiment was used to investigate the effects of feedback
manipulation on decision making of checkers and non-checkers as an intermediate situation compared to the negative-negative experiment.

Additionally, it is important to consider that these pictures presented a real feedback effect in the experiments where they were used as positive (studies 2 and 4) or negative feedback (studies 3, 5 and 6). Regardless of the fact there were no SCR measures (Skin conductance response) or measures of heart rates during the view of these distinct pictures, as a way to measure emotional responses to each picture, the significant effects between experiments in the positive study (neutral pictures vs. positive pictures) and in the negative study (neutral pictures vs. negative) have shown that the addition of positive or negative pictures did influence decision making behaviour in the checker and in the non-checker group (see results at Chapter 6.3.1.2 and 6.3.2.2). Indeed, this effect cannot be explained by money feedback, as money was also present in study 2 (positive-neutral) and study 3 (negative-neutral).

For the lab-based questionnaires, there was no group and experimental effect of impulsivity, while gambling habits and risk-taking varied across studies. Risk-taking varied across experiments, so participants generally risked more in the positive-neutral (lower error magnitude) than in the negative-neutral experiment but they also risked more in the symptom-related (higher error magnitude) than in the negative-neutral experiment (lower error magnitude). In this way, error magnitude cannot explain risk-taking differences across experiments, so it is possible this factor was dependent on the sample used in each study. The same seems to be the case for risk-perception in finance, where one effect was only found in the symptom-related study and for gambling results that varied, depending on the study. One common finding regarding all studies was related to intolerance of uncertainty where checkers were more intolerant to uncertainty than non-checkers. This did not depend on the experiment, so it is possible these findings are more related with a trait usually found in the obsessive compulsive population (even if subclinical) and this corroborates findings that OCD is generally more intolerant to uncertainty than the general population (Hollaway et al., 2006).

Regarding general results for the binary decision making studies, they have shown that both checkers and non-checkers were able to learn the binary RBLT as accuracy increased throughout sub-blocks for all the three blocks. In this way, learning did not depend on the presence of random or non-random conditional probabilities and on the feedback direction, magnitude of the task or the presence of symptom-related emotional stimuli.
Interestingly, Starcke et al. (2010) results for their associative feedback learning task have shown that OCD participants presented an accuracy deficit regarding both their positive and negative experiments, while in our study checkers were able to learn the task, either in a positive-neutral context as in a negative-neutral context. However, it is not possible to compare our results with this study, as their sample was constituted by clinical OCD participants and the experiment was different from ours with about only 32 trials and giving feedback at winnings and losses (no absent feedback). Nonetheless, this indicates that deficits found in decision making of OCD could be also related to the task employed and to the sample characteristics of the study, and this is important to consider because so many experiments use distinct criteria for their samples, as clinical vs. subclinical, discrimination vs. no discrimination between subtypes and presence of medication (see Abramovitch, 2015).

Regarding others analysis in this current thesis, for win-shift it seems groups did not differ in its terms and there was no experiment difference for both groups either. However, there was a marginal difference between the positive-neutral and positive-positive experiment, where participants generally shifted more after winning in the positive-neutral than in the positive-positive experiment. This is possibly the case because the risk associated with the positive-positive experiment is higher as it offers an absence of two rewards after an error. In this way, participants would generally shift more in the positive-neutral experiment. Apart this difference, however, the absence of results regarding win-shift indicate that subclinical checkers do not present a deficit in the use of previous trials after a positive feedback is released in a RBLT.

For cross-correlation in general, participants used the immediately previous trial to make decisions in block 1. This supports the idea that sequential effects are indeed present in random sequences, where participants use previous feedback to make future decisions, even when the target position is not dependent from previous target positions (see Vulkan, 2000; Unturbe & Corominas, 2007; Wilder et al., 2009). However, participants generally do not seem to use other previous trials to make predictions in the positive experiment. Nonetheless, for the negative-negative experiment, participants were also cross-correlating with the second previous trial, what indicates participants were performing more exploration in these experiments. Clearly, if using only one previous trial (positive) or two previous trials (negative), the use of previous trials in block 1 was not dependent on the group, so checker deficits regarding RBLTs seem to not be related with sequential effect deficits that could
affect decision making in random sequences. Actually, results confirmed that checkers were able to learn the pattern structures in blocks 2 and 3, as non-checkers.

For the investigation of the effect of feedback direction on the decision making of subclinical checkers (Chapter 5), results have shown that for block 1, there was no group and experimental effect for accuracy, win-shift, lose-shift and for any of the lags associated with cross-correlation for the positive-neutral and for the negative-neutral experiment in block 1. However, there was a main effect of experiment for block 1, where participants presented a higher accuracy in the negative-neutral experiment in comparison with the positive-neutral experiment. Unfortunately, as these results concern both groups, they are not informative about differences between checkers and non-checkers, so it is not possible to know which group is responsible for this result. However, these results indicate that participants presented a better performance in the negative-neutral experiment, possibly because the risk associated with exploration in this task is higher, so participants preferred to choose the option with most targets (maximisation). For block 3, there was no significant effect of group or experiment for any of these variables. For the strategy questionnaires, there was no effect of group or experiment for both blocks 1 and 3 in the use of a strategy and visualization of a pattern. Actually, significant results concentrated in block 2.

It is noteworthy that checkers presented a higher accuracy in the block with a non-random conditional probability (block 2) than non-checkers and checkers presented a higher accuracy in this block for the negative-neutral experiment in comparison with the positive-neutral experiment. Additionally, checkers were closer to the experimental sequence (Dif) in lags 2 and 4 than non-checkers and checkers were closer to the experimental sequence in these lags in the negative-neutral experiment in comparison to the positive-neutral experiment. Finally, for the strategy questionnaire, there were more checkers who saw a pattern in the negative-neutral experiment than non-checkers and more checkers who saw a pattern in block 2 in the negative-neutral experiment in comparison to the positive-neutral experiment. For non-checkers, however, they were closer to the experimental sequence in lag 5 in the positive-neutral experiment in comparison to the negative-neutral experiment. In this way, these results indicate that while checkers were closer to the pattern in the negative-neutral experiment, non-checkers seemed to be closer to the optimal strategy in the positive-neutral experiment in block 2.

These results confirm the hypothesis that checkers would present a higher performance, more cross-correlations (lower Dif) and see more patterns when exposed to
negative emotional stimuli (negative-neutral experiment). This would be the case because, while exposed to negative stimuli, checkers would present a deficit in using error information to correct their hypothesis about the probabilities controlling the target position in the task. Indeed, it would be expected that errors would be signals of discontinuation (Van Duijvenvoorde et al., 2008) and this would increase exploration or shifts between options (more lose-shift). However, under negative stimuli checkers would not be able to discontinue and explore new hypothesis about their optimal strategy internal model, presenting a bias towards exploitation of old models. These results would supposedly reflect negative prediction error deficits, as negative prediction errors are conceptualized as a signal for correction of the participant’s hypothesis about the task after an unexpected error is released (Maia, 2009).

Interestingly, results for non-checkers support this hypothesis because they were closer to the pattern in the positive-neutral experiment and more distant in the negative-neutral experiment. This indicates that non-checkers are more biased towards exploration when exposed to negative stimuli. Note, however, that non-checkers were not closer to the pattern sequence in lags 2 and 4 in the positive-neutral experiment, as it was the case for checkers in the negative-neutral experiment. Moreover, accuracy and cross-correlation effects for block 2 were also expected for blocks 1 and 3, where checkers were expected to also present a higher accuracy in these blocks. Additionally, it was expected that checkers would be closer to the pattern in block 3 and repeat responses in the option with the highest target probabilities in block 1, increasing cross-correlation for lags 1 to 5. Finally, it was expected that checkers would present less lose-shift in block 1 for the negative experiment and consequently less exploration in this block after an error. Unfortunately, this was not the case. It is possible, however, that results were not significant because of our sample being constituted of subclinical participants. Other possibility is the fact that this study employed a small sample size. So, future studies should address this paradigm with a bigger sample and considering the investigation within a clinical population. In addition, it is possible that exploitation/exploration depends on others factors that were different between blocks, as maximum accuracy (block 1 vs. block 2) or the presence of one or more optimal strategies (block 2 vs. block 3). In this way, new studies should also be designed to furtherly investigate the role of maximum accuracy and number of optimal strategies. In the other hand, however, these results are an indication that checkers possibly present a distinct behaviour when facing
negative feedback opposed to positive feedback and this was confirmed by studies 1 and 2 in Chapter 6 and study 3 in Chapter 7.

In Chapter 6, studies 1 and 2 aimed to investigate the effects of the feedback magnitude manipulation in the decision making of checkers. The positive study (study 2 vs. study 4) focused on the manipulation of the magnitude of positive feedback and study 2 focused on the manipulation of the magnitude of negative feedback. In Chapter 7, study 3 focused on the manipulation of feedback magnitude of negative feedback but this time considering symptom-related stimuli, as one of the negative feedbacks.

If the hypothesis that checkers present a deficit when only facing negative feedback was correct, it would be expected that there would not be any difference between groups and experiments in the positive study in accuracy, lose-shift and cross-correlation. However, when comparing the positive-neutral accuracy, lose-shift and cross-correlation with the positive-positive experiment (distinct feedback magnitudes), it seems that groups are different.

For instance, in the positive study, there was no group and experiment differences for accuracy but there was a lose-shift marginal group difference for the positive-neutral experiment and a significant group difference for the positive-positive experiment, where checkers shifted more after losing in the previous trial than non-checkers in the positive-neutral experiment and shifted less than non-checkers in the positive-positive experiment. Finally, checkers shifted more after losing in the previous trial in the positive-neutral than in the positive-positive experiment. In addition, there was a group difference for cross-correlation for lag 1 in block 1 for the positive-positive experiment, where checkers were closer to the experimental sequence in lag 1 than non-checkers. This goes in accordance with lose-shift findings for the positive-positive experiments, where checkers shifted less after a previous result, so they would consequently cross-correlate less with the previous trial, decreasing the difference between their sequence and the experimental sequence for lag 1. For block 2, there was a main effect of group for lag 3, where checkers were more distant from the original sequence in a general positive experiment than non-checkers, indicating that non-checkers are possibly doing more exploitation in a general positive experiment than checkers. Regarding the strategy questionnaire, there were no group differences for all the three blocks but there were some experiment differences, where checkers saw more patterns in the positive-positive experiment than in the positive-neutral experiment in block 2.
Actually, all these results support the idea that checkers present a deficit in RBLTs that is dependent on the feedback magnitude of the error in the task and this is not related with the direction of the feedback per se but with feedback magnitude, as group and experimental differences were found in the experiments within a positive and negative direction.

In truth, it is interesting that checkers presented a numerically higher probability of shift after losing and consequently more exploration in the experiment with the lowest error feedback magnitude (positive-neutral) but less shifts, less cross-correlations in block 1 and less exploration in the experiment with the highest error feedback magnitude (positive-positive). This is supported by the findings in the strategy questionnaire, where checkers saw more patterns in the positive-positive experiment than in the positive-neutral experiment in block 2 and this indicates that in the experiment with a higher negative feedback magnitude, checkers are less influenced by disruptions in the pattern (more exploitation), as it has happened in the negative-neutral experiment (see above). In this way, this indicates that increasing the feedback magnitude of a task, at least in the positive direction, would increase exploitation in block 1 (less lose-shifts) and would increase the exploitation of the pattern in the pattern sequence of block 2 (more pattern visualization). Note, however, that this last finding was not corroborated by accuracy and cross-correlation results.

Non-checkers, in the other hand, presented the same amount of lose-shifts and cross-correlation in both positive experiments, so their behaviour seem to not vary on the basis of the manipulation of the error magnitude in this study. Additionally, data indicates that non-checkers do less exploration in a general positive experiment, as they were closer to the experimental sequence in lag 3 for a general positive experiment and lag 5 for the positive-neutral experiment. However, this is still unclear because there was no difference between groups for lags 2 and 4 in block 2 that were related with the pattern in this block.

In fact, it is possible that checkers presented a distinct behaviour in the negative-neutral in comparison with the positive-neutral experiment because the error magnitude of this task is lower than the one presented at the negative-neutral experiment. Thus, checkers would present a deficit in experiments with higher error magnitudes closer to the one presented at the negative-neutral experiment. This deficit would be present in the positive-positive experiment either. Clearly, it is complicated to compare error magnitudes between a positive and a negative experiment, as while one presents an absence of reward after a mistake, the other presents a presence of penalty after a mistake. Indeed, there is actually
evidence that the systems modulating the absence of reward and the presence of penalties after a mistake are different (Maia, 2009; Schultz, 2010; Fiorillo, 2013). In this way, the best way to investigate the hypothesis of error magnitude is to compare distinct feedback magnitudes within the same feedback direction, as it was done in Chapters 6 and 7.

Actually, while it is not possible to compare positive to negative directions, results indicate that more shifts and more exploration in the positive-neutral experiment for checkers is exactly the opposite that was interpreted for the accuracy findings in the negative-neutral experiment in Chapter 5. In addition, less shifts and less exploration in the positive-positive experiment is similar to what it was interpreted for the accuracy results of the negative-neutral experiment in Chapter 5. In this way, it could be claimed that checkers are able to correct their errors through negative prediction errors when exposed to error magnitudes as the ones presented in the positive-neutral experiment but not when exposed to error magnitudes that are slightly higher than the ones presented in the positive-positive experiment and so on. If this is correct, checkers would present a bias towards exploitation in any kind of experiment that presented higher error magnitudes than the ones of the positive-positive experiment. Actually, while it is not possible to confirm that the negative-neutral experiment presents a higher error magnitude than the positive-positive experiment, it is possible to compare the negative-neutral error magnitude with others negative experiments with higher error magnitudes. This was done in studies 2 and 3.

For the Negative study, results have shown that checkers presented similar accuracies, lose-shift and cross-correlation for the three blocks between the negative-neutral and the negative-negative experiment. Additionally, checkers saw similar amounts of patterns and used equal amounts of strategies for the three blocks in the strategy questionnaire. This was also true for study 3, so accuracies, lose-shift and cross-correlation, as well the strategy questionnaire results were similar between the negative-neutral and symptom-related experiment for checkers. Non-checkers, in the other hand, shifted more in the negative-negative experiment than in the negative-neutral experiment and equally between the negative-neutral and symptom-related experiment. In addition, they seemed to see more patterns in the negative-negative than in the negative-neutral experiment and more patterns in the symptom-related than in the negative-neutral condition.

So, in this case, while checkers seem to not be able to modulate their behaviour in such error magnitudes as the ones presented in studies 2 and 3, non-checkers seem to be able to modulate their behaviour when the error magnitude increases to such values as the ones
presented at the negative-negative experiment. In truth, this supports the hypothesis that checkers present a deficit in RBLTs, when the error magnitude surpasses a specific error magnitude threshold. It is possible, this threshold is about the level of the error magnitude presented at the negative-neutral experiment. For non-checkers, it seems that they do the opposite, where they are biased towards exploitation in a general positive experiment and biased towards exploration in a general negative experiment. Note that Schultz (2010) argued that the absence of a reward is not followed by a negative prediction error, as it happens after a penalty is released. In this way, it is possible that the presence of exploration in the positive-neutral and exploitation in the positive-positive experiments are indicating others deficits in checkers, not related with the same deficits found for the negative experiments. Indeed, Fiorillo (2013) considers the possibility that two neurotransmitter systems are responsible for absence of feedback and presence of penalty, so this would indicate that two distinct error systems could be in deficit in OCD. New studies should be designed to investigate deficits in OCD when exposed to absence of rewards (errors in positive experiments) and to the presence of penalties (errors in negative experiments).

In relation to group differences, while there was no group effect for accuracy there was a main group difference for lose-shift, where non-checkers shifted more than checkers (more exploration) in a general negative experiment in the Negative study. Interestingly, as shown above, non-checkers were closer to the experimental sequence at lag 3 in block 2 in a general positive experiment, indicating that they do present less exploration in experiments within a positive direction. In this way, while data indicates that non-checkers do more exploration in a general negative experiment, data indicates that in a positive direction, non-checkers do more exploitation. Actually, non-checkers shifted more after losing in the negative-negative experiment, so this also indicates that non-checkers are more explorers in the negative-negative condition than checkers. Finally, the strategy questionnaire data in study 3 also supports this hypothesis, as non-checkers saw more patterns in block 1 than checkers in the symptom-related experiment. Note that in block 1 there are no patterns, so this result indicates that non-checkers are doing more exploration than checkers in block 1 in this condition, as they might be testing more hypothesis and, consequently, searching for more patterns than checkers in this block.

Finally, regardless of the fact checkers presented similar behaviours in the negative study and in the symptom-related study, it would be expected that this group would present a higher accuracy and to be closer to the auto-correlation values of the experimental sequence
in blocks 2 and 3 as well more lose-shifts in block 1 in the three negative experiments and not only in the negative-neutral experiment. In fact, this would corroborate the idea that checkers were indeed biased towards exploitation in such error magnitudes. In fact, while a group difference in accuracy was only found for block 2 in the negative-neutral experiment, data from the symptom-related study has shown that there were two main group differences in the accuracy of blocks 2 and 3, where checkers presented a higher accuracy in these blocks than non-checkers, when comparing both experiments together (negative-neutral and symptom-related). The results for block 2 could be a reflection of the higher accuracy already found for checkers in the negative-neutral experiment. However, the result regarding block 3 corroborates the hypothesis that checkers might possibly be ignoring disruptions in the patterns shown in block 3 in a general negative experiment, increasing performance, what supports the hypothesis that checkers are indeed biased towards exploitation in such environments with higher error magnitudes.

Interestingly, common and significant correlations for checkers were only present in the negative experiments but not in the positive experiments, where supposedly checkers might present a deficit that possibly regards other physiological system, regarding prediction errors (absence of reward – see Fiorillo, 2013). In fact, significant correlations with accuracy were found for intolerance of uncertainty that was always higher in checkers than non-checkers and did not depend on the experiments. Actually, it seems that checker participants with higher intolerance of uncertainty were the ones with the lower performance in block 2 in the negative-neutral and with the higher performance in block 2 for the negative-negative experiment. In addition, they were the ones with a higher performance in block 1 for the symptom-related. Unfortunately, correlations were not common for the same blocks but it indicated that RBLT deficits found in the negative conditions could be causing an increase or a decrease in intolerance of uncertainty of checker participants. The opposite could be also true and intolerance of uncertainty could be one of the causes for the deficits. However, if this was true, intolerance of uncertainty would also affect performance in the positive experiments, where it was found to be higher in checkers too. In this way, it is possible that checker participants have their intolerance of uncertainty increased or decreased depending on the feedback magnitude of the RBLT, when the magnitude surpasses a specific threshold. Note that others correlations were found for each one of the experiments but they were not common to all of them or to a specific reward direction or related to feedback magnitude, so
it is possible they were more related with the sample employed in each study than with a general finding regarding the groups investigated.

In summary, data supports that subclinical checkers without comorbidities present a specific deficit in RBLTs. This deficit seems to not be related with a feedback direction but with values of negative feedback magnitudes that are higher than a specific threshold. Indeed, it seems that the augmentation of the error magnitude of a task influences non-checkers towards exploration, while checkers seem to do the opposite. Actually, this is also corroborated by extra analysis posteriorly performed in this thesis. New analysis were introduced after the end of the study to explore new possibilities, regarding the combination of distinct experiments. Actually, while considering the two dimension hypothesis and separating studies in a positive study (positive-neutral vs. positive-positive experiment) and in a negative study (negative-neutral vs. negative-negative experiment), it was not possible to explore new analysis comparing positive with negative experiments (i.e., negative-negative vs. positive-neutral). For this reason, a Two-way (2 x 5) ANOVA analysis considering all experiments together was posteriorly executed. Secondly, while the symptom-related task is considered an intermediate value of feedback magnitude between the negative-neutral and the negative-negative experiment, there was no analysis considering the comparison between these three experiments together. In this way, a second Two-way (2 x 3) ANOVA considered the comparison between these three negative experiments. Data analysis only considered the gambling task dependent variables, as accuracy, win-shift and lose-shift and cross-correlation. Note that when a significant value was found, post-hoc tests were only executed for experiments because analysis between groups for each experiment were already shown in the previous chapters (see chapters 5, 6 and 7). ANOVA results were resumed using a table that can be seen at Appendix 2.7. Post-hoc significant results are shown and discussed below.

8.2 Extra analysis

8.2.1 Analysis considering all experiments together

A two-way ANOVA between-participants with the factors Group (checkers vs. non-checkers) and Experiment (negative-neutral vs. positive-neutral vs. negative-negative vs. positive-positive vs. symptom-related) was conducted for accuracy in block 1; accuracy in block 2; accuracy in block 3; win-shift and lose-shift, and a three-way mixed ANOVA with the within-participant factor Lag (1-5) and the between-participant factors Group (checkers
vs. non-checkers) and Experiment was computed on the difference values of cross-correlation. As it was previously shown, an effect of lags was always present, as lags were designed to be different, so this effect will not be discussed furthermore. Results can be seen at Appendix 2.7. Basically, there was a significant interaction for accuracy in block 2, $F(4, 108) = 3.27$, $p = .01^*$, $\eta^2 = .10$; a significant interaction for lose-shift, $F(4, 108) = 5.28$, $p < .001^{**}$, $\eta^2 = .16$; a significant interaction for the difference in the correlation values for block 1, $F(16, 432) = 1.80$, $p = .02^*$, $\eta^2 = .06$ and a significant interaction for the difference in the correlation values for block 2, $F(7.44, 432) = 2.68$, $p = .01^*$, $\eta^2 = .09$. Additionally, there was a significant main effect of experiment for lose-shift, $F(4, 108) = 2.76$, $p = .03^*$, $\eta^2 = .09$ and a significant main effect of group for the difference in the correlation values for block 1, $F(4, 432) = 2.98$, $p = .01^*$, $\eta^2 = .02$. All the others analysis were non-significant (see Appendix 2.7).

Basically, post-hoc tests for accuracy in block 2, lose-shift and the Difs in blocks 1 and 2 considering ten comparisons between each of the five experiments, revealed that there was a significant difference between the positive-positive and the negative-neutral experiment in the non-checker group for the Dif (lag 3) in block 2 (Mean (Positive-positive): 0.03 ± 0.06; Mean (Negative-neutral): -0.13 ± 0.04; $t(22)= 3.53$, $p_b < .02$, $r = .60$), where non-checkers were closer to the experimental sequence in the positive-positive in comparison to the negative-neutral experiment. All other comparisons were non-significant for all of the five lags in block 2. In addition, there was a significant difference between the symptom-related and the negative-negative experiment in the non-checker group for lose-shift (Mean (symptom-related): 0.43 ± 0.03; Mean (Negative-negative): 0.61 ± 0.04; $t(22)= 3.41$, $p_b < .03$, $r = .59$), where non-checkers shifted less in the symptom-related experiment than in the negative-negative experiment after losing in the previous trial. Additionally, there was a significant difference between the positive-neutral and the positive-positive experiment in the checker group for lose-shift (Mean (Positive-neutral): 0.56 ± 0.03; Mean (Positive-positive): 0.35 ± 0.05; $t(21)= 3.82$, $p_b < .01$, $r = .64$) that was previously shown in Chapter 6. All other comparisons were non-significant for lose-shift. No other significant effects were found, including analysis for the main effects of group for lose-shift and main effect of experiment for the difference in correlation values for block 1.
8.2.2 Analysis considering all the negative experiments

A two-way ANOVA between-participants with the factors Group (checkers vs. non-checkers) and Experiment (negative-neutral vs. negative-negative vs. symptom-related) was conducted for accuracy in block 1; accuracy in block 2; accuracy in block 3; win-shift and lose-shift, and a three-way mixed ANOVA with the within-participant factor Lag (1-5) and the between-participant factors Group (checkers vs. non-checkers) and Experiment was computed on the difference values of cross-correlation. An effect of lags was always present, as lags were designed to be different, so this effect will not be discussed furthermore. Results can be seen at Appendix 2.7. Basically, there was a significant interaction for accuracy in block 2, $F(2, 66) = 3.60, p = .03^*$, $\eta^2 = .09$ and a significant interaction for the difference in the correlation values for block 2, $F(3.68, 264) = 2.86, p = .03^*$, $\eta^2 = .08$. Additionally, there was a significant main effect of experiment for lose-shift, $F(2, 66) = 3.70, p = .03^*$, $\eta^2 = .09$ (see Appendix 2.7).

Basically, post-hoc tests for accuracy in block 2, Dif in block 2 considering three comparisons between each of the three experiments, did not reveal any significant effect. For the main effects of experiment for lose-shift there was no significant effect after Bonferroni corrections.

8.2.3 Discussion

For analysis 2 and for the others variables, however, there were no significant results at all, when Bonferroni corrections were employed. Considering analysis 1, it was interesting to find that non-checkers were closer to the experimental sequence in the positive-positive experiment than in the negative-neutral experiment. This seems to agree with the hypothesis that the enhancement of the negative feedback magnitude of the task influences non-checkers towards exploration (see Chapter 8.1 for a discussion), when assuming that the error magnitude in the positive-positive experiment is lower than the error magnitude of the negative-neutral experiment. Additionally, this hypothesis is corroborated by the fact that non-checkers shifted less in the symptom-related experiment in comparison with the negative-negative experiment, so again, increasing the error magnitude of the task seems to influence non-checkers towards exploration. In this way, the new findings related to analysis 1 indicate that the hypothesis that non-checkers are biased towards exploration when the error magnitude of the task increases is very plausible and these results fit previous results already
shown in this thesis in Chapters 5, 6 and 7 (see Appendix 2.8 resuming all results). In the other hand, checkers would present the opposite behaviour what could indicate a deficit.

Indeed, It is possible that these deficits are related with negative prediction errors deficits, while deficits regarding errors related with the absence of reward in the positive experiments could indicate others deficits, or related with others neurotransmitter systems responsible for prediction errors (Fiorillo, 2013) or deficits not related with prediction errors. This hypothesis will be discussed in detail section 8.3.

8.3. Deficits in negative prediction errors? Discussion of general hypothesis

The deficit found for checkers in the negative experiments indicate that they are not able to modulate their behaviour in such higher error magnitudes, while they are still able of doing it in lower error magnitudes, as in the positive-neutral experiment. This deficit does not reflect a lower performance per se, especially because results indicate that checkers presented an even higher performance in the negative-neutral experiment than non-checkers for specific blocks. Actually, this deficit is more related with a bias towards exploitation in these tasks, while non-checkers are more prone to perform exploration in such error magnitudes. Indeed, in the particular case of these experiments, this deficit helped to increase performance, so it is important that RBL studies do not focus only on measurements of accuracy when investigating deficits in a specific population.

In summary, however, this thesis have shown that there is a bias towards exploitation in checkers within specific error magnitudes. Other explanation for this bias is that this deficit is not related with specific negative feedback magnitudes but with specific positive feedback magnitudes. This is the case because it is also possible to measure experiments in their amounts of reward magnitudes, where the positive-neutral condition presents the lowest value of reward magnitude in the positive direction and the negative-neutral presents the lowest value in the negative direction (just one absence of penalty). Results, in this case, would indicate that checkers would present a deficit towards higher values of rewards.

Clearly, the first hypothesis related to error magnitude deficits makes more sense because checkers seem to not be affected by the manipulation of reward magnitude (no win-shift effects) and there is evidence that OCD presents a deficit in the generation of negative prediction errors (that compares an unexpected previous result with the current prediction after an error) but not with positive prediction errors. For instance, previous literature has shown that OCD presents lower feedback error related negativity (FRN) than controls (Zhu et
al., 2014) and this ERP seems to reflect negative prediction errors (Holroyd & Coles, 2002) because it is larger after negative feedback (Nieuwenhuis et al., 2004). In this way, if FRNs are lower in OCD, it is possible that negative prediction errors are in deficit in this group after an error. Other important evidence is related to the fact that OCD seems to present a hyperactive dopamine system (Perani et al., 2008). Dopamine seems to be responsible for the generation of both positive and negative prediction errors (Frank et al., 2004; Maia, 2009) but according to Maia (2009), while pauses in the production of dopamine would lead to the generation of negative prediction errors, an increase in the amount of dopamine would produce positive prediction errors. This is very controversial, as errors and rewards could be modulated by distinct dopaminergic or neurotransmitter systems in the brain and not only by pauses or increases in dopamine (Maia, 2009; Schultz, 2010; Fiorillo, 2013). However, based on the first hypothesis that more dopamine would generate positive prediction errors and pauses would generate negative prediction errors, it is assumed that groups with a hyperactivation in the dopaminergic system as OCD would have a lower probability to produce negative prediction errors, as there would be a lower probability for pauses in dopamine production. This would generate deficits in RBLTs because there would not be enough negative prediction errors to correct current hypothesis about the target position, based on unexpected previous results.

Actually, Frank et al. (2004) have shown that unmedicated Parkinson disease (PD) patients (lower dopamine) learned better from negative feedback, while medicated PD patients (high dopamine) learned better from positive feedback. These results indicate that participants with higher levels of dopamine as OCD would learn better from positive feedback and this is similar to our results, where checkers were still able to modulate their behaviour in experiments with positive feedback, while they were not capable to modulate their behaviour in experiments with negative feedback. Note, however that this seems to not depend on the feedback direction (positive vs. negative) but on the error magnitude of the feedback employed in the task, as a bias towards exploitation was already present at a positive experiment (positive-positive). Actually, different from the Frank et al. (2004) study, this thesis investigated more than one error magnitude in a positive direction, as well more than one error magnitude in a negative direction. Thus, here, it was possible to know that the error magnitude seems to be more relevant in the generation of deficits than feedback direction only. As a matter of fact, if the error magnitude is indeed more important than the feedback direction, it would be also possible to find results in which medicated PD patients...
would fail to learn in experiments within positive directions but with higher error magnitudes. New experiments should address this question in the future.

In the other hand, if considering the hypothesis that other neurotransmitter systems could be responsible for negative prediction errors (Fiorillo, 2013), it would make sense that OCD would be affected in their negative prediction error system, as 1) the serotonin system is affected in OCD (Westenber et al., 2007) and it was hypothesized that this system could be responsible for negative prediction errors (Maia, 2009) and 2) because it was hypothesized that specific dopamine receptors could be responsible for negative prediction errors (Maia, 2009; Fiorillo, 2013) and, as the dopamine system is affected in OCD, this would also affect negative prediction errors in OCD.

Clearly, this study presented limitations, so while the hypothesis about a specific deficit with negative prediction errors could explain the results currently found in this thesis, they are only suppositions. Actually there are other explanations regarding the differences between groups. For instance, as this study focused on the investigation of the use of previous trials by participants, it could be argued that working memory deficits could also be reflecting differences between groups, especially because distinct studies have shown the presence of deficits in OCD regarding working memory (see Harkin & Kessler, 2011 for a review). However, if working memory deficits were affecting results, it would be more plausible that checkers would not be able to learn sequential effects in the tasks with patterns (blocks 2 and 3). However, this was not the case and checkers were even better in block 2 than non-checkers in the negative-neutral experiment. Additionally, even if working memory deficits were acting in the results, it would be more plausible that all results would be affected, especially in blocks 2 and 3, regardless of the feedback magnitude and feedback direction of the task. However, results were different depending on the feedback direction and feedback magnitude of the task, so even if working memory deficits were present here, the feedback direction and the feedback magnitude of the task is also influencing results, so the prediction error hypothesis continues to be a possible explanation for the results. Indeed, it is possible that different deficits are acting in the results here, as this task does not exclude the possibility that working memory deficits are influencing data. More studies are needed to investigate possible working memory deficits in subclinical checkers, while doing a RBL task and how working memory deficits are influenced by the manipulation of the feedback magnitude and feedback direction of the task.
In any case, the current thesis results indicate that checkers seem to present a bias towards exploitation in situations where the error magnitude of the negative feedback is above a certain threshold. This would explain why checkers are less biased towards exploration in experiments with higher error magnitudes. Non-checkers, in the other hand, presented the opposite behaviour. In this way, while non-checkers would be biased towards exploration in a particular error magnitude, checkers would be biased towards exploitation. These results need to be discussed on the basis of the limitations of this study. In the next section, the limitations of the current thesis will be discussed, followed by its future directions.

8.3. Limitations

It is recognised that the design and implementation of the studies in this thesis present limitations. First of all, the entire sample employed in our studies is constituted of undergraduate and graduate students. In this way, this sample presents a socioeconomic and educational level that is very characteristic and do not represent the entire community. Other important limitation regards the number of participants presented in each group. Even though this thesis aimed to recruit at least 20 participants per group, the specificity of the checker and non-checker group, regarding the exclusion criteria, decreased the chances of finding enough participants for each group in the amount of time required for the studies. Finally, even though subclinical checkers are widely studied in the literature, these results and interpretation of the main findings cannot be directly transferable to the clinical population.

Actually, it would be quite interesting to study a clinical population and to compare this population with a subclinical one, as a way to investigate the influence of the symptom severity augmentation in the exploitation tendency and consequently, in the participant performance. Indeed, while subclinical checkers might be increasing their performance in block 2 at the negative-neutral experiment, because of their augmented exploitation tendency in specific error magnitudes, it is possible that the symptom severity will potentially increase exploitation to the extreme, preventing participants to employ exploration in the sequence, what is essential for learning pattern and strategies in an uncertain environment. In this way, excessive exploitation, increased by symptom severity would increase performance at one point but would decrease it completely when the symptom severity surpasses a certain threshold. This would create a U-shaped relationship between learning a pattern or a decision making strategy and symptom severity but as this study was developed with subclinical
checkers this assumption is still not clear, so it is important to investigate if symptom severity will negatively affect learning in clinical patients. In truth, this would have clinical implications, as higher or lower symptom severity could indicate learning deficits regarding decision making strategies or even advantages when an uncertain situation shows up.

Regarding the picture validation task, one important limitation of the symptom-related study concerns the fact that checker-related pictures were equally rated by checkers and non-checkers. This could indicate that our checker group was not sensitive enough to the checker-related pictures or that our non-checker group was sensitive enough. Actually, results from the symptom-related study suggested that non-checkers presented a higher amount of depression in comparison to checkers in this specific sample, even though there was no significant difference in depression between groups. In fact, it is possible that the sample employed in this study presented particular characteristics that, not only decreased rating differences for the checker-related pictures but that could also influence results in the binary decision making task. In this way, even though results from the symptom-related study corroborated the main hypothesis of this thesis, it is important to replicate this study with other sample that rate checker-related pictures distinctively between checkers and non-checkers, so results in the binary decision making task could be confirmed.

Regarding the binary decision making study, one of the important limitations of the study concerned the fact that experiments of studies 4, 5 and 6 were recorded in a different university than the ones of the neutral studies 2 and 3. Actually, while the neutral experiment was recorded in Bournemouth University, all the others experiments were recorded at University of Surrey. First of all, it is possible that these two samples presented distinct socioeconomic characteristics but also the rooms where the experiments were recorded were distinct, as well the monitors, where the stimuli were present. Indeed, the measurements for the stimuli in the two universities were slightly distinct, even though this difference was not enough to generate important distinctions in the stimuli.

Considering the experimental design, it is important to highlight that block order was not randomized. Due to the fact that sequential effects could bias participants towards pattern searching and that exposition to the pattern blocks could increase this bias, it was decided to keep the random block as block 1 always. As block 3 also presented block 1 strategy, it was decided to not subsequently expose participants to block 3 after block 1. In this way, block 1 was always first, the block with patterns was always 2 and the block with both strategies was always block 3. Clearly, even if this strategy was justified, the non-randomization of the
block order increased the changes that measurements of blocks 2 and 3 were also related with fatigue or were influenced by previous blocks strategies. So, for next studies, it would be interesting to keep blocks as between-participant variables and not as within-participant ones. Indeed, this was consider as a strategy for the non-randomization of blocks. Nevertheless, if feedback magnitude was consider as within-participant factor, participants would be exposed to one only block strategy five different times and this would constitute other problem.

Additionally, it is important to consider that feedback order was always the same, where money was always shown before and followed by pictures. Actually, this is another limitation of the study, as it is possible that the feedback effect of the pictures are being decreased by the fact that they are being shown after the money feedback. In fact, for future studies, it would be important to counterbalance the feedback order between money and pictures, as a way to control for possible increases or decreases in the feedback effect of pictures or money. However, even though this was not done, it is also important to consider that the pictures did present a feedback effect, as the addition of positive and negative pictures in studies 4 and 5 modified results, when compared to the neutral pictures in studies 2 and 3 (see Chapter 6).

Finally, regarding the prediction error hypothesis, one of the limitations concern the fact that significant differences were not found for all the blocks, so new studies should replicate these experiments, while considering possible reasons for the absence of significant values in specific blocks, as the absence of a clinical population or the number of participants in each group. Finally, one of the biggest limitations of this study is the absence of a prediction error physiological measurement that would correlate with our behavioural data. At first, one EEG study was designed to measure FRNs in checkers and non-checkers, while measuring behavioural data in two distinct feedback conditions. Unfortunately, this study was not concluded due to timing issues. However, it is recognized that measurements of FRNs would be important in future studies to give support to the hypothesis of specific deficits with negative prediction errors in OCD. Indeed, it would be even more interesting if these measurements were done in subclinical and clinical checkers, while performing these RBLTs.

8.4. Future directions

Considering the above synthesis of the thesis findings, there are still a lot of areas where more research is needed. In fact, one of such research concerns the investigation of the effect of others error magnitudes than the ones investigated here on RBL in checkers. It
would be important to consider the investigation of only one feedback direction, especially the positive one, and the manipulation of its error magnitude. This is the case because it is possible that checkers do not present deficits in experiments with error magnitudes that are lower than the ones presented in the positive-neutral study. Additionally, it would be interesting to correlate measurements of exploration and exploitation with physiological correlates of negative prediction errors, as FRNs. Other interesting measure would be the correlation with skin conductance responses (SCRs) and heart rates that are shown to be a correlate of emotional reactivity (Boucsein, 1982). This would increase our knowledge about how behavioural measurements are related with the emotional reactivity of checkers after seeing a feedback.

Importantly, future studies could consider the manipulation of the error magnitude regarding only one type of feedback. Actually, in this study two types of feedback (money plus pictures) were employed due to the need to investigate symptom-related stimuli but new studies could consider the manipulation of only money or picture emotional magnitude values to investigate RBL in OCD, what would decrease the number of factors to be manipulated in the experiment.

Finally, future studies could concentrate on the investigation of the effect of maximum accuracy values in sequences with non-random and random conditional probabilities. In this study, it was not possible to manipulate the maximum accuracy value within the same sequence. However, future studies could address this question, especially because prediction errors are not only related with feedback magnitude but also with feedback probability (Lak et al., 2014) and it is quite possible that checkers also present deficits with distinct feedback probabilities, if they supposedly present deficits with negative prediction errors. In the same way, feedback delay is other area of interest for future studies and results could eventually support the hypothesis of the existence of negative prediction error deficits in OCD.

Additionally, the findings above might be also relevant for OCD treatment. On a rather speculative note, checkers might learn well in positive and negative settings that have a low negative feedback magnitude. However, checkers over-rely on old routines (exploitation) in situations with high negative emotional value, because they might fear that they could not get rewards otherwise. This means that heightened emotions and negative feedback magnitudes might reduce treatment success because of the enhanced exploitation behaviour. In this way, while this suggestion is rather speculative, it would be interesting to perform
further investigations about the effect of feedback magnitude manipulation in stimuli used in cognitive behavioural therapies.

8.5. Conclusions

New studies investigating how the manipulation of feedback magnitude affects OCD are needed to be designed. In fact, these studies, together with other future investigations about the effect of feedback probability and delay, could bring new evidence about specific negative prediction error deficits in OCD and this could help to develop new approaches to treat OCD, either on the therapeutic level, as on the drug level. Actually, the investigation of specific deficits in OCD, regarding prediction errors could bring new leads about how dopamine dysfunctions affect OCD and how this neurotransmitter interacts with serotonin and glutamate, both affected in OCD (Carlsson, 2001). Additionally, this could bring new research developments in the understanding of how negative prediction errors are physiologically correlated, as there is still too much controversy about the sole role of dopamine as its responsible and the existence of negative prediction errors related to the absence of rewards (errors in positive experiments).

From the thesis findings, results indicated that subclinical checkers without diagnosed depression, anxiety, panic disorder and PTSD presented a bias towards exploitation in experiments where the error magnitude surpasses a specific threshold. This bias occurred in both positive and negative directions, so error magnitude seems to be more correlated with magnitude than feedback direction. Interestingly, non-checkers presented a bias towards exploration in such error magnitudes when checkers were more biased towards exploitation.

Indeed, research findings are limited due to the several limitations of this study and because results were not found for all the blocks investigated. In fact, results only indicate a direction of thinking for future studies but it is still possible to explain these findings based on the current literature of OCD, indicating that this population supposedly presents deficits in the generation of negative prediction errors. Actually, it is possible that the findings of this thesis are reflecting deficits in negative prediction errors in OCD, which are a reflection of a hyperactive dopamine system in this population. Future studies should be designed to confirm if this is the case.
References


267


Appendix 2.1 – Copy of ethics approval for University of Surrey studies

University Ethics Committee

Ms Camila Gomes Victorino
School of Psychology
FAHS

08 December 2014

Dear Ms Gomes Victorino

UCE ref: EC/2014/47/FAHS
Study Title: A gambling study looking at decision making in checkers

On behalf of the Ethics Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the submitted protocol and supporting documentation.

Date of confirmation of ethical opinion: 08 December 2014.

The final list of documents reviewed by the Committee is as follows:

<table>
<thead>
<tr>
<th>Document</th>
<th>Version</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Email from supervisor responding to Committee’s outstanding queries in</td>
<td></td>
<td>05 Dec 2014</td>
</tr>
<tr>
<td>email of 05 Dec 2014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Signed letter from researchers responding to Committee’s queries in</td>
<td></td>
<td>03 Dec 2014</td>
</tr>
<tr>
<td>letter of 01 Dec 2014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Signed letter from researchers detailing all changes to resubmitted</td>
<td></td>
<td>19 Nov 2014</td>
</tr>
<tr>
<td>documentation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protocol Cover Sheet</td>
<td>19 Nov 2014</td>
<td></td>
</tr>
<tr>
<td>Phase 1 &amp; 2 Document 4: Summary of the Project</td>
<td>19 Nov 2014</td>
<td></td>
</tr>
<tr>
<td>Phase 1 &amp; 2 Document 5: Detailed Protocol of the Study</td>
<td>19 Nov 2014</td>
<td></td>
</tr>
<tr>
<td>Letter from Collaborator: Johann Briffa</td>
<td>01 Apr 2014</td>
<td></td>
</tr>
<tr>
<td>Phase 1 &amp; 2 Document 8: Risk assessment</td>
<td>19 Nov 2014</td>
<td></td>
</tr>
<tr>
<td>Phase 1 &amp; 2 Document 9: References</td>
<td>19 Nov 2014</td>
<td></td>
</tr>
<tr>
<td>Phase 1 &amp; 2 Document 10: Flow Chart</td>
<td>19 Nov 2014</td>
<td></td>
</tr>
<tr>
<td>Phase 1 Document 1: Recruitment Email- tracked copy</td>
<td>03 Dec 2014</td>
<td></td>
</tr>
<tr>
<td>Phase 1 Document 2: Recruitment Poster</td>
<td>19 Nov 2014</td>
<td></td>
</tr>
<tr>
<td>Phase 1 Document 3: Online Screening Questionnaire- tracked copy</td>
<td>03 Dec 2014</td>
<td></td>
</tr>
<tr>
<td>Phase 1 Document 4: Lab-Based Participant Information Sheet -</td>
<td>03 Dec 2014</td>
<td></td>
</tr>
<tr>
<td>tracked copy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase 1 Document 5: Lab-Based Consent Form</td>
<td>19 Nov 2014</td>
<td></td>
</tr>
<tr>
<td>Phase 1 Document 6: Picture Validation Task Example</td>
<td>19 Nov 2014</td>
<td></td>
</tr>
<tr>
<td>Phase 1 Document 7: Lab-Based Debriefing Document</td>
<td>19 Nov 2014</td>
<td></td>
</tr>
<tr>
<td>Phase 2 Document 1: Recruitment Email- tracked copy</td>
<td>03 Dec 2014</td>
<td></td>
</tr>
<tr>
<td>Phase 2 Document 2: Recruitment Poster</td>
<td>19 Nov 2014</td>
<td></td>
</tr>
<tr>
<td>Phase 2 Document 3: Online Screening Questionnaire</td>
<td>19 Nov 2014</td>
<td></td>
</tr>
<tr>
<td>Phase 2 Document 4a: Lab-Based Participant Information Sheet Exp 1-</td>
<td>03 Dec 2014</td>
<td></td>
</tr>
<tr>
<td>tracked copy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phase 2 Document 4b: Lab-Based Participant Information Sheet Exp 2-</td>
<td>03 Dec 2014</td>
<td></td>
</tr>
<tr>
<td>tracked copy</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 2.2 – Copy of ethics approval for Bournemouth University studies

## Research Ethics Checklist

<table>
<thead>
<tr>
<th>Reference Id</th>
<th>8759</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status</td>
<td>Approved</td>
</tr>
<tr>
<td>Date Approved</td>
<td>12/10/2015</td>
</tr>
</tbody>
</table>

### Researcher Details

<table>
<thead>
<tr>
<th>Name</th>
<th>Ellen Seiss</th>
</tr>
</thead>
<tbody>
<tr>
<td>School</td>
<td>Faculty of Science &amp; Technology</td>
</tr>
<tr>
<td>Status</td>
<td>Staff</td>
</tr>
<tr>
<td>Course</td>
<td>Staff</td>
</tr>
<tr>
<td>Have you received external funding to support this research project?</td>
<td>No</td>
</tr>
</tbody>
</table>

Please list any persons or institutions that you will be conducting joint research with, both internal to BU as well as external collaborators.

- Camila Gomes Victorino (University of Surrey, Bournemouth University)

### Project Details

<table>
<thead>
<tr>
<th>Title</th>
<th>A gambling study looking at decision making in checkers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed Start Date</td>
<td>15/07/2015</td>
</tr>
<tr>
<td>Proposed End Date</td>
<td>30/09/2016</td>
</tr>
</tbody>
</table>

Summary (including detail on background methodology, sample, outcomes, etc.)

- Summary (Document 2) Detailed Protocol (Document 3)
Appendix 2.3 – Example of online questionnaire

A gambling study looking at decision making in checkers

This study received a favourable opinion from the University Ethics Committee of the University of Surrey.

Names and Contact Details of Investigators:

Camila Victorino
Email: c.gomesvictorino@surrey.ac.uk
01483 686870

Dr Ellen Seiss
Email: e.seiss@surrey.ac.uk
01483 686934

What is this about?
We are carrying out research at the University of Surrey, which is investigating the relationship between decision making strategies in people with checking and non-checking characteristics. Checking habits might be the frequent checking of doors, windows, locks, ovens and other objects perceived as possible source of danger. Checkers are often anxious about these objects and frequently think about the objects throughout the day. The objective of this study is to understand decision-making behaviour in checkers and non-checker, as assessed by the performance in gambling tasks.

For this study we are searching for participants who do show checking behaviours at times or who never check and who are willing to rate the emotional influence of the pictures.

At the beginning of the study we will ask you to fill out online screening questionnaires lasting about 15 minutes. To compensate you for your time you will be able to participate in a prize draw where you can win 1 x £30 or 2 x £10 Amazon vouchers.

On the basis of the screening we might invite you to a laboratory-based study where you will be asked to (1) rate positive, neutral, negative, and checking-related pictures, (2) play a gambling task, and (3) fill out some lab-based questionnaires about intolerance to uncertainty, risk-taking behaviour, impulsivity, and your gambling behavior. The lab-based study will take about 1.5 hours and we will pay you £10 for your participation plus what you have won during the gambling task.

We would be grateful if you would help us by taking part in this study. The research could help provide a better understanding of checking behaviours and hopefully assist in the development of future treatments for the clinical Obsessive Compulsive Disorder (OCD) checker population.

What does the study involve?
This study consists of two parts and this online questionnaire is part 1. You can fill out the online questionnaire without committing to part 2 of the study. In this questionnaire, we will ask about past and present symptoms related to mental health, including depression/anxiety and obsessive thoughts and behaviours. The questionnaires should take about 15 minutes to complete. As a thank you, you will be entered into a prize draw for a chance to win an Amazon voucher (1 x £30 or 2 x £10). This prize draw is independent of your participation in the second part of the study. It is important to note that this study is a mere assessment of the presence or absence of checking-related behaviour. This behaviour is rather common in the general population (~10 - 15%) and it is not equivalent to a diagnosis of OCD.

Are there any downsides of taking part?
You may find some of the questions in the questionnaires quite personal. However, this is needed to find out if it is suitable to invite you to participate in the second part of the study. However, all the information you provide is strictly confidential and would be stored in a password-secured computer. Please feel free to contact us if these instructions are not clear or you would like more information. Besides, if you are concerned about the questions or your answers to these questionnaires please contact us. There are no additional benefits and no risks involved in this study.

How do I agree to take part?
If you decide to take part you will be asked to tick the boxes below. Please, tick the box stating you are in good health and know no reason why you should not participate.

[CONTINUE BUTTON]

Please, tick the box stating you are willing to participate. The next screen will display the online consent form.

[CONTINUE BUTTON]
Participant consent form

Please, tick all the boxes, stating you read each paragraph of the consent form

[TICK BOX] I have read and understood the previous Information Sheet provided. I have been given a full explanation by the investigators of the nature, purpose, location and likely duration of the study, and of what I will be expected to do. I have been advised about any discomfort which may result. I have been given the opportunity to ask questions on all aspects of the study and have understood the advice and information given previously. I have been given the opportunity to provide the information I am in good health and I know no reason why I should not participate. I confirm I continue to be in good health.

[TICK BOX] I expect to be willing and able to comply with the requirements of the study that have been outlined to me, with any instruction given to me during the study and to co-operate fully with the investigators. I shall inform them immediately if I suffer any discomfort, or experience any unexpected or unusual symptoms.

[TICK BOX] I understand that all personal data relating to volunteers is held and processed in the strictest confidence, and in accordance with the Data Protection Act (1998). I agree that I will not seek to restrict the use of the results of the study on the understanding that my anonymity is preserved.

[TICK BOX] I understand that I am free to withdraw from the study at any time without needing to justify my decision and without prejudice.

[TICK BOX] I acknowledge that in consideration for completing the study I shall possibly receive some prize money won in the online and lab-based experiment. I recognize that the sum would be less, and at the discretion of the Principal Investigator, if I withdraw before the completion of the study.

[TICK BOX] I understand that in consideration of completing the study, undiagnosed or untreated issues, related to anxiety, depression, obsessive compulsive disorder and gambling can be found. I am aware that these questionnaire measures do not give a diagnosis related to these issues and are a mere measure of individual differences in behaviour. I recognize that support measures will be offered to me by the research team at the end of my participation, if concern occurs.

[TICK BOX] I confirm that I have read and understood the above and freely consent for participating in this study. I have been given adequate time to consider my participation and I expect to be willing and able to comply with the instructions and restrictions of the study.

Please sign the consent form by ticking this box       [TICK BOX]
# Questionnaire 1

The information you provide in this questionnaire, will be treated as strictly confidential.

**Project supervisor:** Ellen Seiss  
**PhD student:** Camila Gomes Victorino

<table>
<thead>
<tr>
<th>Question</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>[ANSWER BOX]</td>
</tr>
<tr>
<td>Gender</td>
<td>F [TICK BOX] M [TICK BOX]</td>
</tr>
<tr>
<td>Do you play videogames or computer games?</td>
<td>YES [TICK BOX] NO [TICK BOX]</td>
</tr>
<tr>
<td>If yes, what kinds (i.e. virtual casino games; fighting, shooting, fantasy, real life games etc)?</td>
<td>[ANSWER BOX]</td>
</tr>
<tr>
<td>How many hours do you play a day?</td>
<td>[ANSWER BOX]</td>
</tr>
<tr>
<td>Do you play for money?</td>
<td>YES [TICK BOX] NO [TICK BOX]</td>
</tr>
<tr>
<td>If yes, how often?</td>
<td></td>
</tr>
<tr>
<td>Do you gamble?</td>
<td>YES [TICK BOX] NO [TICK BOX]</td>
</tr>
<tr>
<td>If yes, how often?</td>
<td></td>
</tr>
<tr>
<td>Do you use glasses or corrective contact lenses?</td>
<td>YES [TICK BOX] NO [TICK BOX]</td>
</tr>
<tr>
<td>Have you ever had been diagnosed with:</td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>YES [TICK BOX] NO [TICK BOX]</td>
</tr>
<tr>
<td>Anxiety disorder</td>
<td>YES [TICK BOX] NO [TICK BOX]</td>
</tr>
<tr>
<td>Panic Disorder</td>
<td>YES [TICK BOX] NO [TICK BOX]</td>
</tr>
<tr>
<td>Post-traumatic Stress disorder</td>
<td>YES [TICK BOX] NO [TICK BOX]</td>
</tr>
<tr>
<td>Do you currently suffer from (diagnosed or undiagnosed):</td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>YES [TICK BOX] NO [TICK BOX]</td>
</tr>
<tr>
<td>Anxiety disorder</td>
<td>YES [TICK BOX] NO [TICK BOX]</td>
</tr>
<tr>
<td>Panic Disorder</td>
<td>YES [TICK BOX] NO [TICK BOX]</td>
</tr>
<tr>
<td>Post-traumatic Stress disorder</td>
<td>YES [TICK BOX] NO [TICK BOX]</td>
</tr>
</tbody>
</table>
Questionnaire 2

The following statements refer to experiences that many people have in their every day lives. Please click the number that best describes HOW MUCH that experience has DISTRESSED or BOTHERED you in the PAST MONTH.

The numbers refer to the following verbal labels:
0 = Not at All
1= A Little
2= Moderately
3= A lot
4= Extremely

1. I have saved up so many things that they get in the way. 0 1 2 3 4
2. I check things more often than necessary. 0 1 2 3 4
3. I get upset if objects are not arranged properly. 0 1 2 3 4
4. I feel compelled to count while I am doing things. 0 1 2 3 4
5. I find it difficult to touch an object when I know it has been touched by strangers or certain people. 0 1 2 3 4
6. I find it difficult to control my own thoughts. 0 1 2 3 4
7. I collect things I don’t need. 0 1 2 3 4
8. I repeatedly check doors, windows, drawers, etc. 0 1 2 3 4
9. I get upset if others change the way I have arranged things. 0 1 2 3 4
10. I feel I have to repeat certain numbers. 0 1 2 3 4
11. I sometimes have to wash or clean myself simply because I feel contaminated. 0 1 2 3 4
12. I am upset by unpleasant thoughts that come into my mind against my will. 0 1 2 3 4
13. I avoid throwing things away because I am afraid I might need them later. 0 1 2 3 4
14. I repeatedly check gas and water taps and light switches after turning them off. 0 1 2 3 4
15. I need things to be arranged in a particular order. 0 1 2 3 4
16. I feel that there are good and bad numbers. 0 1 2 3 4
17. I wash my hands more often and longer than necessary. 0 1 2 3 4
18. I frequently get nasty thoughts and have difficulty in getting rid of them. 0 1 2 3 4
# Questionnaire 3

Please read each statement and circle a number 0, 1, 2 or 3 which indicates how much the statement applied to you over the past week. There are no right or wrong answers. Do not spend too much time on any statement.

The rating scale is as follows:

- 0 Did not apply to me at all
- 1 Applied to me to some degree, or some of the time
- 2 Applied to me to a considerable degree, or a good part of time
- 3 Applied to me very much, or most of the time

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I found myself getting upset by quite trivial things</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>I was aware of dryness of my mouth</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>I couldn’t seem to experience any positive feeling at all</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>I experienced breathing difficulty (e.g., excessively rapid breathing, breathlessness in the absence of physical exertion)</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>I just couldn’t seem to get going</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>I tended to over-react to situations</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>I had a feeling of shakiness (e.g., legs going to give way)</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>I found it difficult to relax</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>I found myself in situations that made me so anxious I was most relieved when they ended</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>I felt that I had nothing to look forward to</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>I found myself getting upset rather easily</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>I felt that I was using a lot of nervous energy</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>I felt sad and depressed</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>I found myself getting impatient when I was delayed in any way (e.g., lifts, traffic lights, being kept waiting)</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>I had a feeling of faintness</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>16</td>
<td>I felt that I had lost interest in just about everything</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>17</td>
<td>I felt I wasn’t worth much as a person</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>18</td>
<td>I felt that I was rather touchy</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>19</td>
<td>I perspired noticeably (e.g., hands sweaty) in the absence of high temperatures or physical exertion</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>I felt scared without any good reason</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>21</td>
<td>I felt that life wasn’t worthwhile</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>
### Reminder of rating scale:

0  Did not apply to me at all  
1  Applied to me to some degree, or some of the time  
2  Applied to me to a considerable degree, or a good part of time  
3  Applied to me very much, or most of the time

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>I found it hard to wind down</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>I had difficulty in swallowing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>I couldn’t seem to get any enjoyment out of the things I did</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>I was aware of the action of my heart in the absence of physical exertion (e.g., sense of heart rate increase, heart missing a beat)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>I felt down-hearted and blue</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>I found that I was very irritable</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>I felt I was close to panic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>I found it hard to calm down after something upset me</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>I feared that I would be &quot;thrown&quot; by some trivial but unfamiliar task</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>I was unable to become enthusiastic about anything</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>I found it difficult to tolerate interruptions to what I was doing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>I was in a state of nervous tension</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>I felt I was pretty worthless</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>I was intolerant of anything that kept me from getting on with what I was doing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>I felt terrified</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>I could see nothing in the future to be hopeful about</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>I felt that life was meaningless</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>I found myself getting agitated</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>I was worried about situations in which I might panic and make a fool of myself</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>I experienced trembling (e.g., in the hands)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>I found it difficult to work up the initiative to do things</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Prize draw

If you wish to take part in the Amazon voucher price draw, please tick here [TICK BOX]

Please, write your e-mail address, so that we can contact you after the draw [ANSWER BOX]

Please write your first name for e-mail contact purposes [ANSWER BOX]

Invitation
…to participate in the 2nd part of the study

You may be invited to take part in the second part of the study. This involves the completion of some questionnaires, a picture rating task, and a gambling task in the psychology lab. All of this will take approximately 1.5 hours of your time. You will be compensated for your time with a payment of £10 plus some gambling money.

In order to be able to contact you about your possible participation in the study we will ask you to provide us with your email address and indicate on the online survey if you would be willing to take part in the second part of the study.

Participants will be eligible based on their questionnaire answers. It is important to note that being eligible for this study is not an indication of Obsessive-Compulsive Disorder or checking personality. If you are not eligible for the second part, you will be informed by e-mail.

Your e-mail address will be kept confidential and it can only be accessed by the researchers in the Brain and Behaviour Group. Please be aware that you are still free to withdraw at any time and without explanation.

I would like to take part in the 2nd part of this study [TICK BOX]

My email address is: [ANSWER BOX]

Please note that you also need to provide your email address if you want to take part in the prize draw. This prize draw is independent of the second study. Thank you!

Thank you for your time and interest.
Appendix 2.4 – Example of participant information sheet (PIS)

A gambling study looking at decision making in checkers

This study has received a favourable ethical opinion from the University of Surrey Ethics Committee.

Participant information sheet

Version 1 (November 19, 2014)
(experiment 2)

What is this about?

We would like to invite you to participate in a research project at the University of Surrey examining the strategies in people with checking behaviour compared to non-checkers, when they are gambling money. Checking habits might be the frequent checking of doors, windows, locks, ovens and other objects perceived as possible source of danger. Checkers are often anxious about these objects and frequently think about the objects throughout the day. This behaviour is rather common in the general population (approx. 10-15%) and it is not equivalent to a diagnosis of OCD, as there are strong variations of this behaviour across individuals.

Do I have to take part?

It is entirely up to you. Before you decide whether you want to take part, it is important for you to read the following information carefully. Feel free to discuss it with others if you wish. You can also get in contact with us via email or telephone (details below) if you would like more information.

What does the study involve?

If you decide to take part in the study you will be asked to come to the university for a lab-based study that lasts about 90 minutes. You will receive £10 as compensation for your time when taking part in this study. This participation money is guaranteed to you.

This study has 3 parts. First, you will be asked to rate some emotional (positive, neutral, negative, checking-related) pictures in a picture validation task lasting about 15 minutes.

Second, you will take part in a computer-based gambling task which involves choosing between two playing cards which appear on a computer screen. The aim is to find the hidden prize (upright money bag) when you select one of the two cards and to not select the incorrect card. We will give you some start up gambling money at the beginning of the experiment (£10). Your task is to keep as much money as possible. All gambling money you have kept at the end of this experiment will be yours. For each correct card choice, you will not win or lose any money and you will see a neutral picture. For each incorrect card choice, you will lose 2 pence and see a moderately negative picture. The gambling task will last about 40-50 minutes (with breaks).

Finally, you need to complete some questionnaires about your handedness, intolerance of uncertainty, impulsivity, risk-taking behaviour, gambling behaviour, and a colour test.

What are the risks and disadvantages of taking part?

There are no known risks about taking part in the study and studies where similar pictures have been used have not been shown to cause any distress. You are entitled to end your participation at any stage during the study without giving the researcher an explanation. In the event that your participation does cause you distress, you will be advised to seek help from trained professionals from the university wellbeing centre (Tel: (0)1483 68 9498).

Who can take part in the study?
We are looking to recruit people aged between 18 and 32 years old, who are in good health and have no self-reported history of psychiatric or neurological illness.

**What happens if something goes wrong?**

The University of Surrey holds insurance which covers claims for injury or deterioration in health which arises directly from participation in the study but it applies only in those situations where the University can be shown to be legally liable.

**How long will the study take?**

The study will take approximately 1 hour and 30 minutes. You will be compensated for your time and effort with a payment of £10 for participation, and the prize money you win in the virtual card game.

**Can I take breaks?**

Yes. The study is run in three parts; i.e. a picture validation task, a gambling task, and lab-based questionnaires. You can rest between these tasks but there will be also additional breaks in between.

**Will my participation be kept confidential?**

If you decide that you would like to be a participant, you will be asked to sign a form giving your informed consent to take part. All other data will be anonymised. You are also free to withdraw yourself and your data at any time in the process without having to give the researchers any explanation. Your participation in this study will be completely anonymous. All information collected will be coded and stored according to an allocated participant number to ensure confidentiality. All data files will be treated with the utmost confidentiality. No individual will be named or identified in any report or possible publication arising from this study. Data will be handled in accordance with the Data Protection Act 1998 and kept for 10 years.

**Does the experiment have therapeutic benefits?**

No. Please note that this project is for research purposes only. It does not aim to provide practical therapeutic benefits or information.

**If you have any questions, or are unsure about any aspect of the experiment, please contact us in advance to talk about your possible participation.**

Thank you
Camila Gomes Victorino
Dr Ellen Seiss

**Further information**

If you want to find out more about the study, please contact Camila Victorino or Dr Ellen Seiss using the contact information below.

**Complaints or concerns**

Any complaints or concerns about any aspects of the way you have been dealt with during the course of the study will be addressed. Please contact Dr Ellen Seiss using the information below.

Dr Ellen Seiss  
[e.seiss@surrey.ac.uk](mailto:e.seiss@surrey.ac.uk)  
01483 686934

Camila Victorino  
[c.gomesvictorino@surrey.ac.uk](mailto:c.gomesvictorino@surrey.ac.uk)  
01483 686870
Appendix 2.5 – Consent forms

A gambling study looking at decision making in checkers

Participant consent form

Please, tick all the boxes, stating you read each paragraph of the consent form

☐ I the undersigned voluntarily agree to take part in the study.
☐ I have read and understood the Information Sheet provided (version 1, November 18, 2014). I have been given a full explanation by the investigators of the nature, purpose, location and likely duration of the study, and of what I will be expected to do. I have been advised about any discomfort which may result. I have been given the opportunity to ask questions on all aspects of the study and have understood the advice and information given as a result.
☐ I expect to be willing and able to comply with the requirements of the study that have been outlined to me. I shall inform the research team immediately if I suffer any discomfort, or experience any unexpected or unusual symptoms.
☐ I understand that all personal data relating to volunteers is held and processed in the strictest confidence, and in accordance with the Data Protection Act (1998). I agree that I will not seek to restrict the use of the results of the study on the understanding that my anonymity is preserved.
☐ I understand that I am free to withdraw from the study at any time without needing to justify my decision and without prejudice.
☐ I acknowledge that in consideration for completing the study I shall receive the sum of £10 plus some money won during the gambling task. I recognize that the sum would be less, and at the discretion of the Principal Investigator, if I withdraw before the completion of the study.
☐ I understand that in consideration of completing the study, undiagnosed or untreated issues, related to anxiety, depression, obsessive compulsive disorder and gambling can be found. I am aware that these questionnaire measures do not give a diagnosis related to these issues and are a mere measure of individual differences in behaviour. I recognize that support measures will be offered to me by the research team at the end of my participation, if concern occurs.
☐ I confirm that I have read and understood the above and freely give my consent to participate in this study. I have been given adequate time to consider my participation and I agree to be willing and able to comply with the instructions and restrictions of the study.

Name of volunteer (BLOCK CAPITALS).............................................................
Signed..........................................................
Date...........................................................

Name of researcher/person taking consent (BLOCK CAPITALS) ............................
Signed..........................................................
Date...........................................................
**Full title of project:** A gambling study looking at decision making in checkers

**Name, position and contact details of researcher:**
Dr Ellen Seiss (room: P330; eseiss@bournemouth.ac.uk, phone: 01202 962373)
Camila Gomes Victorino (room P104; cgomesvictorino@bournemouth.ac.uk)

**Name, position and contact details of supervisor (if the researcher is a student):**
Dr Ellen Seiss (room: P330; eseiss@bournemouth.ac.uk, phone: 01202 962373)

**Name of Contact Details of Senior Staff Member in case of complaints:**
Prof Dr Matt Bentley (email: mbentley@bournemouth.ac.uk)

<table>
<thead>
<tr>
<th>Please Initial Here</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

I confirm that I have read and understood the participant information sheet for the above research project and have had the opportunity to ask questions.

I understand that my participation is voluntary and that I am free to withdraw up to the point where the data is anonymised, without giving reason and without there being any negative consequences. In addition, should I not wish to answer any particular question(s), complete a test or give a sample, I am free to decline.

I give permission for members of the research team to have access to my anonymised responses. I understand that my name will not be linked with the research materials, and I will not be identified or identifiable in the report or reports that result from the research.

I agree to take part in the above research project.

<table>
<thead>
<tr>
<th>______________________</th>
<th>_______________</th>
<th>__________________________________</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name of Participant</td>
<td>Date</td>
<td>Signature</td>
</tr>
<tr>
<td>______________________</td>
<td>_______________</td>
<td>__________________________________</td>
</tr>
<tr>
<td>Name of Researcher</td>
<td>Date</td>
<td>Signature</td>
</tr>
</tbody>
</table>

Once this has been signed by all parties the participant should receive a copy of the signed and dated participant consent form, the participant information sheet and any other written information provided to the participants. A copy of the signed and dated consent form should be kept with the project’s main documents which must be kept in a secure location.
### STRATEGY QUESTIONNAIRES

PUT PARTICIPANT CODE HERE: ____________

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>If yes, which one?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did you use any strategy to find the target in BLOCK 1?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ ] Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ ] No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did you see any patterns in BLOCK 1?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ ] Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ ] No</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>If yes, which one?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did you use any strategy to find the target in BLOCK 2?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ ] Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ ] No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did you see any patterns in BLOCK 2?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ ] Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ ] No</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>If yes, which one?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did you use any strategy to find the target in BLOCK 3?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ ] Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ ] No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did you see any patterns in BLOCK 3?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ ] Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ ] No</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 2.7 – Statistical table of extra analysis, considering all experiments together (analysis 1) and all negative experiments (analysis 2)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Analysis 1 (All experiments together)</th>
<th>Analysis 2 (All negative experiments)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Two-way ANOVA</td>
<td></td>
</tr>
<tr>
<td>Accuracy Block 1</td>
<td>$F(4, 108) = 1.88, p = .11$</td>
<td>$F(2, 66) = 0.40, p = .67$</td>
</tr>
<tr>
<td></td>
<td>$F(1, 108) = 0.71, p = .39$</td>
<td>$F(2, 66) = 0.68, p = .41$</td>
</tr>
<tr>
<td>Accuracy Block 2</td>
<td>$F(4, 108) = 0.03, p = .86$</td>
<td>$F(2, 66) = 0.25, p = .67$</td>
</tr>
<tr>
<td>Accuracy Block 3</td>
<td>$F(4, 108) = 1.41, p = .23$</td>
<td>$F(2, 66) = 1.56, p = .21$</td>
</tr>
<tr>
<td>Win-shift</td>
<td>$F(4, 108) = 1.15, p = .33$</td>
<td>$F(2, 66) = 0.72, p = .48$</td>
</tr>
<tr>
<td>Lose-shift</td>
<td>$F(4, 108) = 2.76, p = .03*, \eta^2_{\text{group}} = .09$</td>
<td>$F(2, 66) = 3.70, p = .03*, \eta^2_{\text{group}} = .16$</td>
</tr>
<tr>
<td>Mixed ANOVA</td>
<td>$F(16, 432) = 2.98, p = .01*$</td>
<td>$F(2, 64) = 1.43, p = .09$</td>
</tr>
</tbody>
</table>

|                    | Mixed ANOVA                                                                                           |                                               |
|                    | $F(16, 432) = 1.80, p = .02*, \eta^2_{\text{group}} = .06$                                             |                                               |
|                    | $F(4, 264) = 0.93, p = .48$                                                                           |                                               |

<p>|                    | Mixed ANOVA                                                                                           |                                               |
|                    | $F(8, 264) = 1.78, p = .09$                                                                           |                                               |
|                    | $F(4, 264) = 1.43, p = .22$                                                                           |                                               |</p>
<table>
<thead>
<tr>
<th>Dif block 2</th>
<th>F(7.44, 432) = 1.43, p = .18</th>
<th>F(1.86, 432) = 1.02, p = .35</th>
<th>F(7.44, 432) = 2.68, p = .01*</th>
<th>F(3.68, 264) = 1.33, p = .26, ηp²_group = .09</th>
<th>F(1.84, 264) = 1.40, p = .24</th>
<th>F(3.68, 264) = 2.86, p = .03*</th>
<th>ηp²_group = .08</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dif block 3</td>
<td>F(9.61, 432) = 0.93, p = .50</td>
<td>F(2.40, 432) = 1.13, p = .33</td>
<td>F(9.61, 432) = 1.68, p = .08</td>
<td>F(4.74, 264) = 1.07, p = .37</td>
<td>F(2.37, 264) = 2.66, p = .06</td>
<td>F(4.74, 264) = 1.30, p = .33</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 2.8 – Table resuming all significant results between groups and between experiments for accuracy in blocks 1, 2 and 3, win-shift and lose-shift and Dif of correlation values for blocks 1, 2 and 3

<table>
<thead>
<tr>
<th>Variables</th>
<th>Between groups</th>
<th>Between experiments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive-neutral</td>
<td>Positive-positive</td>
</tr>
<tr>
<td>Accuracy block 1</td>
<td>similar</td>
<td>Similar</td>
</tr>
<tr>
<td>Accuracy block 2</td>
<td>similar</td>
<td>Similar</td>
</tr>
<tr>
<td>Accuracy block 3</td>
<td>similar</td>
<td>Similar</td>
</tr>
<tr>
<td>Win-shift</td>
<td>similar</td>
<td>Similar</td>
</tr>
<tr>
<td>Lose-shift</td>
<td>similar</td>
<td>similar</td>
</tr>
<tr>
<td>Dif block 1</td>
<td>similar</td>
<td>similar</td>
</tr>
<tr>
<td>Dif block 2</td>
<td>similar</td>
<td>similar</td>
</tr>
<tr>
<td>Dif block 3</td>
<td>similar</td>
<td>similar</td>
</tr>
</tbody>
</table>
