Learning in affective contexts

The impact of incidental anxiety and outcome valence on decision-making and confidence

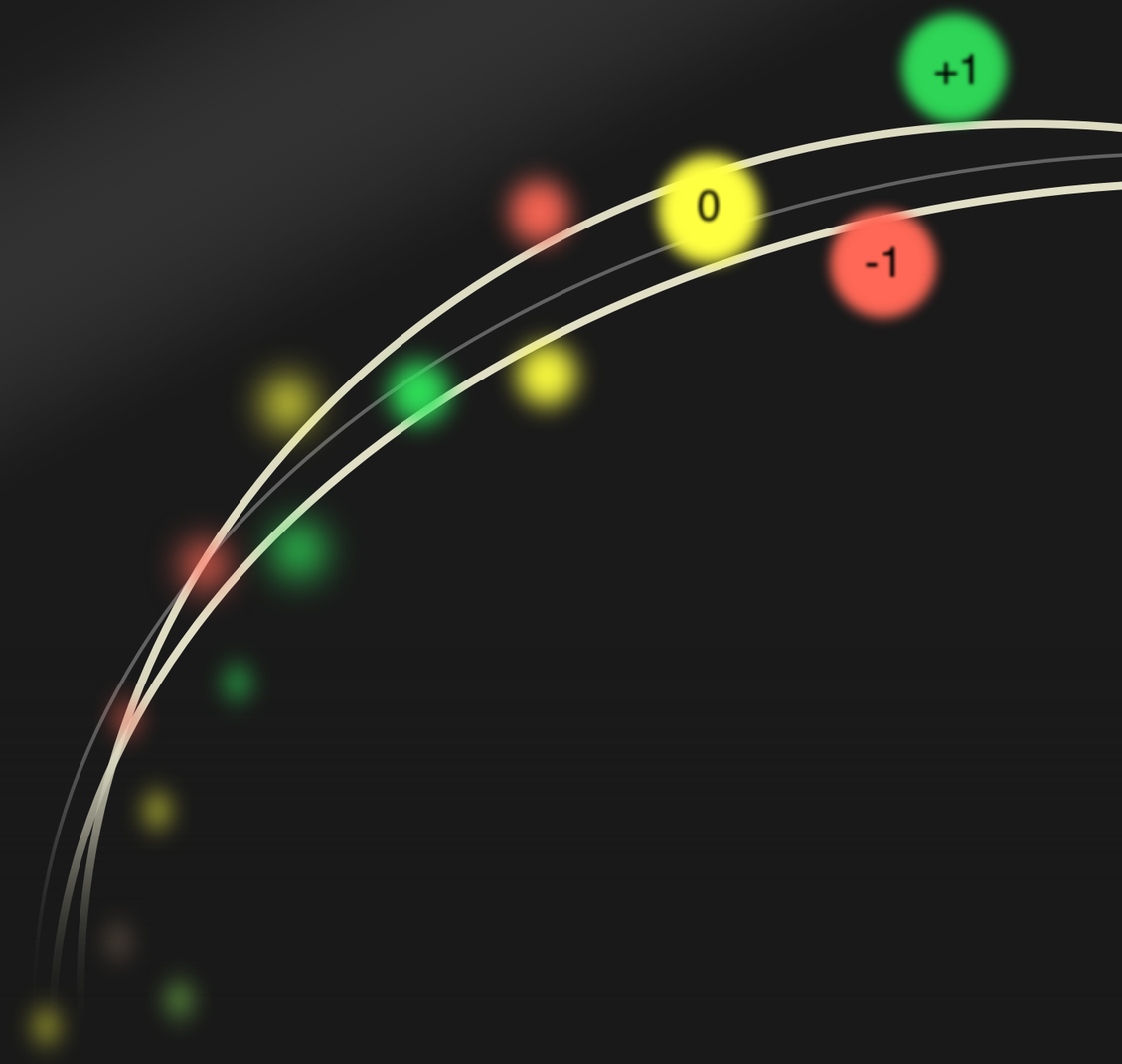
Ting, C.-C.

Publication date
2021

Document Version
Final published version

Learning in Affective Contexts:

The Impact of Incidental Anxiety and Outcome Valence on Decision-Making and Confidence



CHIH-CHUNG TING

Learning in Affective Contexts

CHIH-CHUNG TING

Learning in Affective Contexts

The Impact of Incidental Anxiety and Outcome Valence on Decision-Making and Confidence

Chih-Chung Ting

ISBN: 978-94-6421-208-2

Cover Design: Chih-Chung Ting

Printing: IPSKAMP PRINTING

Learning in Affective Contexts
The Impact of Incidental Anxiety and Outcome Valence on Decision-Making and Confidence

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op maandag 22 februari 2021, te 14.00 uur

door Chih-Chung Ting
geboren te Taichung City, Taiwan

Promotiecommissie

Promotor:	dr. J.B. Engelmann	Universiteit van Amsterdam
Copromotor:	dr. M.P. Lebreton	Université de Genève
Overige leden:	prof. dr. J.H. Sonnemans	Universiteit van Amsterdam
	prof. dr. S. Shalvi	Universiteit van Amsterdam
	dr. R.J. van Holst	AMC-UvA
	prof. dr. C.K.W. de Dreu	Universiteit Leiden
	dr. J. Hausfeld	Universiteit van Amsterdam
	prof. dr. F.A.A.M. van Winden	Universiteit van Amsterdam

Faculteit Economie en Bedrijfskunde

Contents

CHAPTER 1

Introduction ***1***

1.1 Reinforcement Learning (RL) 6

1.2 Metacognition 10

1.3 Making Decision in Affective Contexts 11

1.4 Incidental Anxiety and Decision-Making 15

1.5 Outline of the Thesis 24

CHAPTER 2

The Effect of Anxiety on Reinforcement Learning ***27***

2.1 Introduction 28

2.2 Material and Methods 32

2.3 Results 44

2.4 Discussion 56

CHAPTER 3

Robust Valence-Induced Biases on Motor Response and Confidence in Human Reinforcement Learning ***63***

3.1 Introduction 64

3.2 Material and Methods 66

3.3 Results 72

3.4 Discussion 79

CHAPTER 4

The Neural Mechanisms of Confidence in Reward Learning and Punishment Learning ***83***

4.1 Introduction 84

4.2 Material and Methods 86

4.3 Results 96

4.4 Discussion 103

Summary ***107***

Appendices ***111***

APPENDIX A: Supplementary Material for Chapter 2 112

A.1 Small-scale Literature Review 112

A.2 Demographics and behavior results 115

A.3 Robustness Tests: Effects of experiment 124

A.4 Robustness tests: Effects of trait anxiety 126

A.5 Robustness Tests: Assessment of the Cognitive Resources Hypothesis 128

A.6 Robustness Tests: Selectively test high SCR-responders 130

A.7 Robustness Tests: Performance in early and late learning stages 134

A.8 Robustness tests: Association between 'late' learning stage and preference 136

APPENDIX B: Supplementary Material for Chapter 3 138

B.1 Overview of Experimental paradigms 138

B.2 Overview of treatment effect 140

B.3 Results of GLME 141

B.4 Effect of Experiments 145

B.5 Robust regression analyses 147

APPENDIX C: Supplementary Material for Chapter 4 148

C.1 Demographics and behavior results (with model-free analyses) 148

C.2 fMRI analysis and results 152

C.3 Instruction 155

Bibliography ***162***

Netherlandse Samenvatting ***177***

Acknowledgements ***179***

CHAPTER 1

Introduction

How consistent are human decisions and preferences across different contexts, especially ones that impact our affective states? The recent Covid-19 pandemic has demonstrated that preferences can easily be shaped by contexts that pose a threat to human life and likely trigger strong emotional reactions in most people (Lep et al., 2020; see also Cheng & Cheung, 2005 for the impact of other epidemic). A recent article in the Economist for instance clearly showed that demand for certain consumer products, including flour, ice cream and alcohol, increased significantly during the pandemic ("The pressure points in Britain's supply chain" The Economist, 18.08.2020). These real-world observations accord with prospect theory and empirical evidence from behavioral economics and neuroeconomics showing context-dependent subjective utility (Engelmann & Hein, 2013; Plassmann et al., 2008; Tymula & Plassmann, 2016). Moreover, similar observations have been made at the neural level, such that activation patterns in response to the same choice option can be modulated by context (Elliott et al., 2008; Engelmann & Hare, 2018; Louie et al., 2015; Phelps et al., 2014)[1]. A large body of research thus supports the idea that decision-making can easily be shaped by contexts, which can impact affective states (Cohn et al., 2015; Hsee & Rottenstreich, 2004; Kamstra et al., 2003; Lerner et al., 2015; Loewenstein, 2000; Meier et al., 2019; Peters et al., 2006; Pulcu & Browning, 2019; Rottenstreich & Hsee, 2001) and cognitive processes supporting decision-making, such as learning, memory and attention (Engelmann & Hein, 2013; Erev & Roth, 1998; King-Casas, 2005). However, the impact of context on choice behavior is not consistent, which raises a question: whether and how affect is involved in preference formation and belief updating. To address this issue, behavioral economists and neuroeconomists seek to understand the psychological and neurobiological mechanisms behind decision-making. This is also important to better understand intra-/inter-individual differences in preferences and belief updating within the same context: why do some people show increased demand for certain consumer products during affective contexts, such as the Covid-19 pandemic? What are the affective and cognitive processes that are influenced by such contexts, and how do these differ across individuals?

In the last decades, the theoretical and practical tools from various disciplines have been brought together to investigate the mechanisms of decision-making. One influential model that summarizes these efforts is the model of adaptive decision-making proposed by Rangel et al., (2008) (**Figure 1.1A**). Specifically, Rangel and his colleagues (2008) reviewed studies from economics, psychology and computational neuroscience and proposed that decision can be decomposed into five basic stages of computational processes: (1) the *representation* of the choice problem, internal states (e.g., hungry, happy, anxious) and environments; (2) *valuation* through assigning and computing costs and benefits for the feasible options given the representation of context; (3) the *action selection* and execution given the preceding value comparison; (4) after feedback, *evaluating outcomes* from the chosen (as well as unchosen) options and; (5) *learning* and updating beliefs about context, option values and actions based on received outcome(s) to adapt future choices (**Figure 1.1A**). Because each computational stage captures a specific information process required by decision-making, Rangel's model provides a framework to investigate, as well as to interpret "how" choice patterns

[1] The influence of context on valuation reflects the physical limits of action potentials in neurons. An action potential is the language neurons use to convey information, simplistically speaking like the ones and zeros in computer language. The rate of action potentials is not infinite but ranges from 0 – ca. 200 spikes per second. Therefore, the limited rate of action potentials cannot encode unlimited option values in the world. In order to deal with this biological limitation, neurons adapt the range of action potential to a given context, which is termed range adaptation.

are influenced by contexts. For instance, individual differences in food preferences might be attributed to either the representation of the options (e.g., a focus on health labels rather than sustainability labels) or the valuation of the options (e.g., foods with high-sugar content are discounted more than foods that come at high costs to the environment) or both. The understanding about which stage in decision-making is impacted by specific factors is crucial to offer mechanistic explanations of decision-making that can help in intervention design, as well as policy formulation.

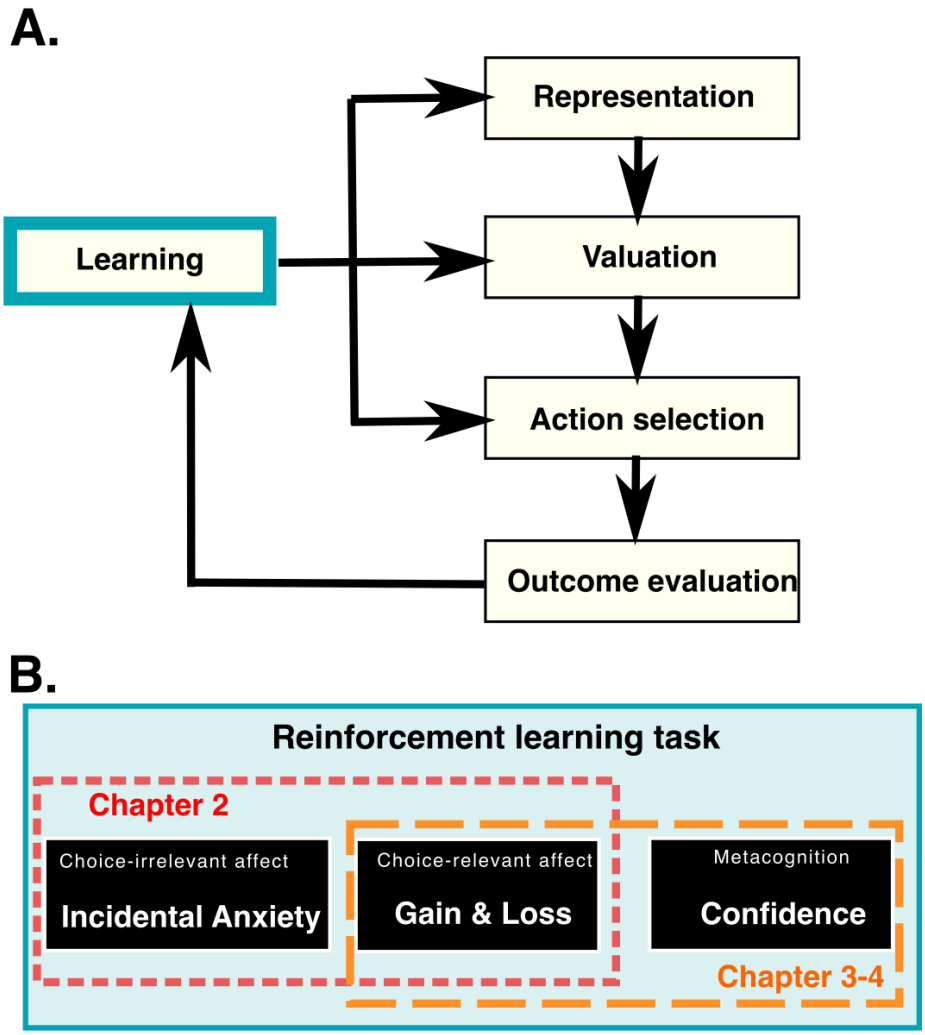


Figure 1.1. Five computational processes of decision-making and the overview of the present thesis. (A) Five stages of computational processes involve in adaptive decision-making. In the present thesis, I mainly focus on the learning stage, which is circled with bold line. The figure is adapted from Rangel et al., (2008). **(B)** The schematic overview of series experiments in the present thesis. Basically, the learning tasks with different context manipulations (i.e., Incidental anxiety and outcome valence) and measurements are used in three empirical studies (**Chapters 2-4**) to answer specific research questions. Red dotted rectangle refers to chapter 2: the effect of incidental anxiety on reward and punishment learning. Orange dashed rectangle refers to chapters 3-4: the effect of outcome valence on confidence judgments.

The current thesis focuses on the role of affect in the learning stage of the decision cascade outlined in **Figure 1.1B**, because the learning stage is relevant to the question mentioned above: "whether and how affect is involved in preference formation and belief updating". Moreover, while recent evidence suggests that affective states can impact choice-related neural circuitry and interfere with decision-making (Barrett et al., 2007; Engelmann et al., 2015, 2019; Robinson et al., 2013; Talmi et al., 2009; Reviews: Engelmann & Hare, 2018; Phelps et al., 2014), little is known to date about how learning is processed under specific affective contexts. Studying the impact of affective states on learning requires experience-based tasks, in which participants learn about the consequences of their decisions via trial and error. A task that offers such an experimental environment is the reinforcement learning (RL) task[2] (rather than one-shot task or any task assuming trials are independent), which we adapted in our studies (**Figure 1.1B**). The trial-dependent feature of RL tasks reflects the fact that choice preferences are modulated by the feedback participants receive for their choice throughout the task, i.e., an option becomes more attractive if previous feedback was positive on average (typically via monetary gains), and negative if previous feedback was negative on average (typically via monetary losses). In addition to the learning task, reinforcement learning (RL) models are able to capture these trial-to-trial dependencies. That is, instead of using averaged choice pattern to identify the impact of context, RL models offer sufficient flexibility and can be fitted to learning data obtained in wide-ranging contexts and tasks, as parameters can be added based on the theoretical information processing during learning and parameters are dynamically adjusted given the performance history and experimental factors (see **Section 1.1 Reinforcement learning** for details and examples).

In order to fill the gap in the literature, we investigate two specific sources of affective information, choice-irrelevant emotion (incidental emotion) and choice *per se*. Specifically, we frame the learning context with choice-irrelevant anxiety induced by uncontrollable and unpredictable electric shock and with outcome valence (i.e., gain vs loss) manipulations (**Figure 1.1B**). First, anxiety is one of the most common emotions in our daily lives and is likely to be carried over from the choice-irrelevant events to the decision problems we are facing. Previous research and resulting theories have speculated that anxiety alters the functioning of the cognitive processes that support decision making, such as attention, memory and information integration (Engelmann & Hare, 2018; Grupe, 2017; Robinson, Vytal, et al., 2013). This hypothesis is supported by neural results showing that the anatomical and functional neural circuits for valuation overlap with anxiety-related neural circuitry (Grupe, 2017; Hartley & Phelps, 2012). However, such empirical evidence is merely suggestive. Moreover, it remains unclear if the impact of anxiety is valence-dependent, i.e., whether anxiety specifically influences decisions over negative outcomes, such as losses, or whether it also influences positive outcomes in the domain of gains. This lack in basic research can lead to difficulties in developing evidence-based treatments for individuals with affective disorders (Grillon et al., 2019). For example, if anxiety selectively influences learning when anticipated feedback (or outcomes) are negative, then treatment developments should take into account such valence-dependent effect and focus on negative information processing. Second, most studies investigating the impact of outcome valence on decision-making focus on choice patterns and

[2] Participants in the instrumental reinforcement learning task are asked to select the option between two fixed-paired options repeatedly. The feedback for the chosen option (with/without the feedback for the unchosen option) are displayed in the end of each trial. In this case, the preference for each option will be changed given outcome (i.e., performance history) (see **Section 1.1.1** for details).

reaction time, the impact of outcome valence on confidence judgments is sparse to date. Confidence judgments are used to index metacognition, which is the ability to track and evaluate one's own status (e.g., health status, work status, driving skill). Deficits in confidence judgments, for instance overconfidence (i.e., the expected status is higher than realistic status), can negatively impact economic and health outcomes for the decision-maker (Berner & Graber, 2008; Camerer & Lovallo, 1999; Malmendier & Tate, 2005). Moreover, a recent review and study further suggested that a confidence bias might be an important indicator of mental illness because psychiatric symptoms are consistently associated with increased/decreased confidence and under-/over-confidence rather than task performance (Hoven et al., 2019 and Rouault et al., 2018). Despite the important role of confidence in decision-making and the fact that the interaction between anticipated gain/loss and confidence has received much attention recently (Giardini et al., 2008; Lebreton et al., 2018), its mechanism is still unclear. In order to systematically investigate value- and confidence-updating in these two specific affective contexts, I review the literature investigating effects of anxiety and outcome valence on value-based decision-making[3] to identify shortcomings in experimental and theoretical approaches. Accordingly, I improve experimental approaches for the experiments reported in subsequent empirical studies and place results in the context of both model-free and model-based approaches to better understand learning process in different affective contexts.

This thesis consists of three empirical studies (**Chapters 2-4**) using a reinforcement learning task with two affective framing manipulations (incidental anxiety and outcome valence) to address the research questions: Whether and how incidental anxiety and outcome valence impact learning process? And, what the underlying affective, cognitive, metacognitive and neural processes are that explain this impact. Specifically, subjects are asked to recognize the best symbol via trial and error when the environment is framed with choice-irrelevant emotion, namely induced anxiety (**Chapter 2**), as well as when the choice is framed as gains or losses (**Chapters 2-4**). In chapters 3 and 4, we focus on the effect of outcome valence and also assess the effects of outcome valence on metacognitive processes in the form of confidence judgements.

Before continuing to the empirical studies reported in Chapters 2-4, the sections below will first review the existing literature about three terms: Reinforcement learning, metacognition and two specific affective framing manipulations, that are relevant to the topics (**Figure 1.1B**). The structure for the rest of this introduction is described as follows. In the first subsection, I introduce reinforcement learning and its applications, including the potential benefits of taking "performance history" into account when studying the impact of emotion on decision making. In the second subsection, I will review the current knowledge about the role of metacognition in decision-making. Next, I will introduce the current understanding of outcome valence and anxiety and their impacts on valuation and decision-making.

[3] Besides value-based decision-making, another type of decision in neuroeconomics is "perceptual decision-making". Unlike value-based decisions, the motivation is relatively meaningless in perceptual decision-making and the answers are usually predetermined given objective and physical properties of options (e.g., whether the size of two circles are the same). This is not the case for value-based decisions because the correctness in value-based decision is usually determined by assumptions (e.g., utility, equilibrium states).

1.1 Reinforcement Learning (RL)

The ability to update beliefs and adjust decisions given experience is crucial for individuals to adapt to new and changing environments and, ultimately, to increase their chances of survival. Considering the fact that choices in real-life are not totally independent of experiences, it is important to implement a framework to account for experiences to explain decisions and the formation of underlying preference. Reinforcement learning (RL) is a well-documented algorithm quantifying trial-and-error learning and belief updating (Rescorla & Wagner, 1972; Sutton & Barto, 1998). Like many existing models, the RL approach offers a mathematical framework to quantify choice patterns given environment. Moreover, these quantified values enable researchers to measure/analyze the experimental contexts (Bavard et al., 2018; Lebreton et al., 2019; Palminteri et al., 2015) and to link behavioral data with physiological responses (e.g., brain activity, pupil dilation and eye movement). These are crucial features to look into the "mechanisms" supporting decision formation and to explain individual differences in decision-making. Importantly, the algorithm is supported by neuroscientific findings showing that dopaminergic neurons (in ventral tegmental area: VTA) track anticipated outcomes and their firing rates reflect a fundamental computational process of reinforcement learning models: the prediction error (Schultz et al., 1997). Prediction errors capture the discrepancy between the anticipated and actual outcome, and are commonly referred to as learning signals.

The RL framework provides a way of understanding how preferences are formed under different contexts and how decisions are influenced by experiences. Here, I summarize at least three advantages of using RL models to study decision-making:

(1) **RL models can identify hidden functions that support value-updating**. The common approach of using aggregated performance (e.g., averaged choice frequency from each subject) in learning tasks only reveals limited information about the underlying learning process, which is a dynamic process that changes over time and depends on the individual. RL models address this limitation by fitting data trial-by-trial and by optimizing relevant parameters, including the learning rate that reflects the speed of acquiring new information. These parameters (e.g., learning rate) are usually associated with latent variables (e.g., Q-values and Prediction errors) that are not detectable from averaged performance and reflect the variable of interest. This feature not only provides a way of estimating underlying mechanisms of information processing, but also provides flexibility for researchers to add parameters for variable of interests, for instance to model different (affective) contexts.

(2) **RL models take into account performance history.** The latent variables in RL models are affected by the history of each individual's choices, experienced outcomes and other psychological factors. This feature allows RL models to flexibly describe and predict choice patterns over time given previous choices (Erev & Roth, 1998; Lee et al., 2012; van den Bos et al., 2013).

(3) **Linking behavioral data to physiological data.** The expected option value and prediction errors are estimated trial-by-trial. In the other words, these values are dynamically updated. This feature enables researchers to test correlations between estimations and physiological data gathered from, for instance, functional magnetic resonance imaging (fMRI) and electroencephalography (EEG).

Given these advantages, the reinforcement learning framework has been broadly implemented in fields of economic (Erev & Roth, 1998, 2014; Roth & Erev, 1995; van den Bos et al., 2013), computational psychiatry (Chen et al., 2015; Frank et al., 2004; Huys et al., 2020; Rutledge & Adams, 2017) and neuroeconomics (Konovalov & Krajbich, 2019) to better understand the mechanisms underlying observed behavioral patterns and preference formation. For instance, van den Bos et al., 2013 found that bidding in the five-player value auction (i.e., subjects bid for the option displayed on the screen and highest bid was determined as winner) over time can be well-captured by an RL model that includes social parameters (i.e., feeling of win and feeling of loss in the auction) and outperforms predictions based on the Nash equilibrium. That is, winning/losing an auction on a preceding trial makes individual bid with a higher/lower value, respectively, on the next trial. These social factors were supported by fMRI results as two common social cognitive brain regions, temporoparietal junction (TPJ) and anterior insula (AI), were associated with social win (i.e., win the auction) and social loss (i.e., loss the auction), respectively. This example clearly demonstrates three benefits of using RL models mentioned above in a practical way: (1) RL models can identify hidden functions that support value-updating; (2) Parameters are dynamically affected; (3) Linking behavioral data to physiological data.

1.1.1 Instrumental Learning Task and Preference

In the context of *instrumental* reinforcement learning, participants are incentivized to identify the best option among multiple choice options (i.e., the one that enjoys highest expected value) via trial-and-error learning. Instead of passively experiencing symbol-outcome associations (i.e., Pavlovian conditioning), instrumental learning tasks require participants' engagement by making decisions within the same set of symbols (usually two symbols) which are associated with different probabilistic outcomes. In order to maximize long-term profits, participants make decision and accumulate outcomes from the chosen symbol to either maximize rewards or minimize losses.

Previous studies have used different methods to assess participants' learning performance: either directly during the acquisition of new information, which is noisy at the beginning and then asymptotes toward the end of the experiment, or by using an extinction test (i.e., making decision without feedback anymore) after learning is completed. In the learning period, learning performance is determined as the percentage of choosing the symbol with the higher expected value in each pair. The history of learning can be observed via a learning curve (see **Chapter 2: Figure 2.4A** for an example), which summarizes the average number of subjects choosing the better symbol at each trial. Unlike the learning task measuring learning *per se* or the amount of value-updating given the outcome history, the performance in the post-learning task (aka transfer task) is used to assess how subjects evaluate the learned symbols (see **Chapter 2: Figure 2.5A** for an example).

1.1.2 Traditional learning model and its Extensions

The decisions made in the instrumental learning tasks have been successfully modeled using various reinforcement learning models[4] (Montague et al., 2004; Sutton & Barto, 1998). Typically, these models work as follows: when a decision maker is facing a novel environment that offers potential rewards or potential punishments, initial decisions are assumed to be random, but the decision-maker can learn from the feedback provided by the environment. Decision strategies are adapted based on a *prediction error signal*, which occurs when the received outcome (say that attainment of some reward, or avoidance of some punishment) deviates from expectation, and *learning rate*, which scales linearly with the magnitude of this deviation. The traditional Q-learning model quantifies this procedure via two essential elements mentioned above: prediction error (PE) and learning rate. The prediction error refers to differences between expected option value (Q) and actual outcome (R), which can be expressed in the following notation:

$$PE_t = R_t - Q_t$$

The prediction error may not be fully implemented and updating expected values is a gradual process. The expected value of the decision for the next trial (t) therefore depends on both a prediction error and a weight applied to the prediction error, modeled as the *learning rate* that reflects the decision-maker's efficiency of utilizing the prediction error as follows:

$$Q_{t+1} = Q_t + \alpha \times PE_t$$

A higher learning rate indicates that the individual is sensitive to the current outcome and tends to use feedback from the prediction error signal to update option values. A higher learning rate is required for fast value-updating in highly volatile environments. In contrast, in relatively stable environments a lower learning rate is preferred, although behavioral adaptation may be relatively inefficient.

After updating the option value, it will not be directly transferred to binary choice. Instead, the expected option values are used to calculate the probability of choosing option *A* over option *B* in the context of a sigmoid function, which is illustrated by the following softmax rule:

$$P(A) = ((1 + \exp\left(Q(B) - Q(A)\right)) \times \beta)^{-1}$$

The temperature parameter β is commonly included in RL models to capture the tradeoff between exploration and exploitation. When a decision is exploration-oriented, the value difference between (Q(A) and Q(B)) is no longer informative as decision is made with less consideration of Q(A) and Q(B). In this case, the model applies a lower β parameter, which reduces the reliance on discrepancy between two options and forms flattened sigmoid function to predict the choice. In contrast, exploration-oriented decisions rely heavily on value differences between choice options, and a higher β enhances the influence of value differences on prediction of the final choice[5]. These

[4] Reinforcement learning models are currently drivers of artificial intelligence development (Barto & Sutton, 1997)

[5] Palminteri and Pessiglione (2017) used figures to illustrate how α and β modulate learning curves and learning performance.

parameters have been used to identify how emotion alters learning processes (Browning et al., 2015; Chen et al., 2015; Robinson, Overstreet, et al., 2013). For instance, the negative correlation between learning rate and level of trait anxiety (Browning et al., 2015) suggested that anxiety not only modulates the neural processes of the prediction error (Robinson, et al., 2013), but also alters the speed of value updating in volatile environments.

The impacts of experimental manipulations on decisions made in the learning tasks have been successfully modeled using various reinforcement learning models. That is, the traditional RL model (Q-learning) has been used and extended by adding parameters to identify how specific effects of interest affect learning. For example, the effects of depression (Chen et al., 2015), learning context (Palminteri et al., 2015), racial bias (Lindström et al., 2014) as well as additional cognitive biases, such as the confirmation bias (Palminteri, Lefebvre, et al., 2017) and optimism bias (Garrett & Daw, 2020; Lefebvre et al., 2017) have been successfully modeled in past research. For instance, Palminteri and his colleagues (2015) demonstrated that not only option value, but also context value associated with specific combinations of choice options is updated via trial and error. The context value is supposed to become positive if a set of options are consistently associated with positive events (e.g., monetary gains in reward learning). On the contrary, context value is updated and becomes negative if options are associated with negative events (e.g., monetary losses in the punishment learning). Within the same study, they further showed that the learned context values are used to adjust outcome evaluation. Specifically, omitted rewards are evaluated as a bad outcome in the pure-gain condition and omitted losses are evaluated as good outcomes in the pure-loss condition (Palminteri et al., 2015; Seymour & McClure, 2008). This example clearly reflects the notion that context value (or the reference point) is dynamically changed and is involved in outcome evaluation (Elliott et al., 2008; Kahneman, 2011; Kahneman & Tversky, 2000; Seymour et al., 2015). However, updating context value is not necessary to be included in the learning process for every participant. For instance, Palminteri et al., (2016) demonstrated that adolescents and adults applied different learning strategies to the same learning task, as adolescents' decisions was best captured by a simpler Q-learning model, while adults' decisions relied on a more complex RL model that included parameters for context value updating. Together, this reflects the high flexibility of Q-learning models to capture both experimental manipulations and specific features in subgroup of participants that can affect learning mechanisms.

Yet, while adding parameters based on specific hypotheses is possible and important, it remains equally important to assess proper model fit as well as to perform parameter and model recoverability assessments. Unlike assessments of goodness-of-fit, the parameter and model recoverability is used to ensure that the winning model (and its parameters) can generate a specific choice patterns (Palminteri, Wyart, et al., 2017). This can be done by simulating data based on models of interest first, and subsequently test if simulated datasets can be correctly attribute to the model used to generate the data (i.e., model recoverability) and if estimated parameters correlated with simulated parameters (i.e., parameter recoverability) (see Palminteri, Wyart, et al., 2017 for the step-by-step practical applications). This is even important for researchers using parameters to represent particular behavioral phenotype (i.e., risk aversion, loss aversion). The good model/parameter recoverability indicates the parameter enjoys unique concepts and consequently, strengthen the link between parameter and behavioral pattern of interest. Given these features of

RL models, the present thesis carefully adapted and evaluated RL models to capture learning strategies in different contexts.

1.2 Metacognition

Decision-making has been extensively studied in terms of choice patterns and reaction times, but research on the ability to evaluate and to regulate one's thoughts and decisions, so called metacognition, is sparse to date. In the decision-making field, metacognitive processes (e.g., self-evaluation of actions) are commonly defined and measured as the belief of making correct choices among multiple options (Fleming & Dolan, 2012; Yeung & Summerfield, 2012; Meyniel et al., 2015; Pouget et al., 2016). In particular, Pouget et al., (2016) suggested that "*defining confidence as the probability that the current choice, overt or covert, is correct.* …". This claim is consistent with the notion that decisions are typically accompanied by a confidence judgment (De Martino et al., 2013; Lebreton et al, 2015; Folke et al., 2016; Shapiro & Grafton, 2020). Moreover, a recent neural evidence suggests that confidence judgments arise automatically after a decision (Shapiro & Grafton, 2020). Therefore, not only confidence judgments but confidence accuracy that is defined as the discrepancy between confidence and overall probability of being correct, are usually used to quantify metacognitive function.

Maintaining accurate confidence judgments is crucial to balance prior and incoming information (Meyniel et al., 2015). This is because confidence, and metacognitive processes in general, are important to flexibly change decision strategies and adapt to changing environments (Donoso et al., 2014; Heilbron & Meyniel, 2019; Vinckier et al., 2016). For example, individuals with high confidence about their predictions about who will become the next president might be less likely to change their prediction and tend to ignore news conflicting with their ideas. From this example, it is easy to see how overconfidence can lead to decision biases and why such biases have been receiving such attention recently (Kahneman, 2011). Accordingly, confidence biases (e.g., over-confidence: the confidence is higher than actual probability of being correct on average) may highly correlate with cognitive biases in information integration, including the status-quo bias (i.e., assigning more weights to the prior information than incoming information to form the decision) and the base-rate fallacy (i.e., assigning more weights to the incoming information than prior information to form the decision) as they reflect the unequal weights on prior knowledge and current information. As referenced in Kahneman's Thinking fast and slow (2011) – overconfidence not only impacts individuals' decisions, but can also lead to detrimental economic losses (Malmendier & Tate, 2005; Camerer & Lovallo, 1999)[6] and adverse health outcomes (Berner & Graber, 2008). In addition to the impact of confidence bias on decision-making, the importance of quantifying metacognitive function by assessing confidence and confidence accuracy has recently been proposed as a new and important transdiagnostic approach for assessing mental

[6] Professional investors and CEOs tend to overestimate their success rates, which results in excessive trading. This overconfidence occurs despite the fact that these populations have abundant information about stock markets.

health (Hoven et al., 2019; Rouault et al., 2018). Rouault et al. (2018) found that symptoms of psychiatric disorders were consistently associated with alterations on metacognitive function rather than cognitive performance. While these findings and reviews pointed out that metacognition plays an important role in various fields, the understanding about emotion-cognition-metacognition interaction and its mechanisms in the decision-making field is inadequate.

To better understand the mechanisms of confidence and the impact of experimental factors on confidence, potential difficulties in investigating confidence in decision-making should be emphasized. Confidence is usually associated with task performance and reaction time. Specifically, better performance and shorter reaction times have been linked to higher confidence (Fontanesi et al., 2019; Maniscalco & Lau, 2012) as they might represent the same concept: task difficulty (van den Berg et al., 2016; De Martino et al., 2013; Moran et al., 2015; Pleskac and Busemeyer, 2010; Ratcliff and Starns, 2009, 2013; Yu et al., 2015). Therefore, it is crucial to dissociate these confounding factors from confidence judgments using proper experimental design, so that interpretations about the source of confidence biases in decision-making can be identified.

1.3 Making Decision in Affective Contexts

To investigate how decision is influenced by a certain affective state, one straightforward way is to compare the performances between the affective and the neutral contexts. The affective context can be artificially created by inducing choice-irrelevant emotion (e.g., emotional responses to Covid-19) on healthy participants or by manipulating the choice-relevant characteristics, like outcome valence (i.e., gains vs losses). The present thesis selectively investigates choice-irrelevant anxiety (i.e., incidental anxiety) and choice-relevant affect, namely outcome valence. Anxiety and outcome valence are commonly existed in real-life decisions and are suggested to guide/bias decisions (Lerner et al., 2015). However, incidental anxiety and outcome valence may influence decision-making differently. Specifically, incidental anxiety is independent of the goal of decision-making and is likely to alter decisions based on the nature of anxiety-related responses (e.g., avoidance, sensitive to negative events). By contrast, outcome valence directly determines the motivation of decision as individuals have tendency to approach gains and avoid losses. In order to understand how affective states of interest (i.e., incidental anxiety and outcome valence) influence decision-making, I will briefly introduce current understanding of the impact of outcome valence on valuation in this section, and introduce incidental anxiety and decision-making in the next section (**Section 1.4 Anxiety and decision making**).

1.3.1 Outcome valence and decision-making

Outcome valence (i.e., gains versus losses) carries affective information as it induces opposing subjective feelings and guide choices by approaching positive affect and avoiding negative affect. The impact of outcome valence on decision-making has been widely investigated and the results

showed that choice patterns are different between gain-framed and loss-framed contexts. For instance, when choosing between risky and safe options (with the same expected value), the risky one is less preferred when the safe option is described as "keeping" a certain amount from the endowment (e.g., keep $60 from $100 for sure). In contrast, the risky option is preferred when the safe option is described as losing a certain amount from the endowment (i.e., lose $40 from $100 for sure), for the same sure outcome (Kahneman & Tversky, 2000). This example demonstrates that the decision problem is strongly influenced by external contextual factors, such as the description of choices, and the choice pattern is changed accordingly, even when the consequences of the decisions are the same. This phenomenon is called valence-framing effect, which has been extensively used to investigate the main motivation of decisions, like dishonest behavior (i.e., lie for avoiding losses or lie for gaining rewards; Abe & Greene, 2014; Schindler & Pfattheicher, 2017) and donation behaviors (Das et al., 2008).

The difference on choice patterns between gain-framed and loss-framed contexts might be the results of the asymmetrical scale for gains and losses (Kahneman & Tversky, 1979). This unbalance processing is illustrated as loss aversion. Monetary outcome is commonly used as proxy to quantify the proportion of weighting of losses and gains. One of classic task used to measure loss aversion is the mixed gambles task, in which subjects are required to select between two options, one gain and one loss, with the same probability (i.e., 50%) (De Martino et al., 2010; Tom et al., 2007). Another common task requiring subjects to select between two options, one mixed gamble (50% gain and 50% loss) and one guaranteed amount of zero, has been used to assess both loss aversion and risk aversion simultaneously (Charpentier et al., 2016; Engelmann et al., 2015). In either task, the gain and loss amounts are varied throughout the experiment and orthogonalized, so that choices can be fitted via a random utility model (i.e., logistic regression) that yields a loss aversion parameter (Sokol-Hessner et al., 2009; Sokol-Hessner & Rutledge, 2019; Tom et al., 2007). Higher loss aversion indicates higher sensitivity to losses compared to equivalent gains.

At the neural level, gains and losses are processed in the same neural circuits, including prefrontal cortex (i.e., Orbital frontal cortex: OFC; and ventral prefrontal cortex: vmPFC) and striatal brain area (i.e., caudate, putamen and nucleus accumbens), which are usually associated with motivation and decision-making. Nevertheless, these regions process gains and losses with different ways as enhanced activation is usually associated with gains and deactivation is usually associated with losses (Bartra et al., 2013; Tom et al., 2007). However, the brain area processing losses-related information is still controversial (Oldham et al., 2018; Seymour et al., 2015). Some studies suggested that losses processing requires additional neural circuits, including anterior insula (AI) and amygdala (Engelmann et al., 2015, 2017; Palminteri, et al., 2012; Cartoni et al., 2016; Holmes et al., 2010; Talmi et al., 2008). One possible reason for the controversial results is that the subjective gains and losses might be changed by other decision-related information, like alternative outcomes and previous performance. Specifically, the objective losses might be reframed as gain when loss is absent or when the alternative outcome is worse. Following the same logic, omission of gains might be reframed as losses (Elliott et al., 2008; Nieuwenhuis et al., 2005; Seymour et al., 2015).

1.3.2 Outcome valence in reinforcement learning

Previous studies also showed that outcome valence can influence cognitive processes during the learning stage, which combines the previous and present information about choice options and further update information processing (**Figure 1.1A** and **Section 1.1 Reinforcement learning**). In the context of reward learning, subjects are incentivized to maximize benefits by choosing those options that are associated with higher rewards on average. By contrast, in the context of punishment learning subjects learn to avoid negative outcomes by choosing those options that are associated with lower losses on average. While many studies have demonstrated that reward and punishment can be equally learned when establishing action-outcome associations in reinforcement learning tasks, some studies showed that the relatively good option learnt from loss contexts is preferred over the relatively bad option learnt from gain contexts (Klein et al., 2017; Lebreton et al., 2019; Palminteri et al., 2015). The findings implied that the lower gains in the reward learning and the lower losses in the punishment learning are reframed as bad and good, respectively, despite the fact that a lower gain is equal to or better than a lower loss. This phenomenon has been explained by recently developed reinforcement learning models. Specifically, the option value is updated based on both performance history and "context value", which is usually computed as average of options' values (Bavard et al., 2018; Palminteri et al., 2015).

Besides distinct motivations for reward and punishment learning, Palminteri and Pessiglione (2017) suggested that distinct neural mechanisms are required for gain-seeking and loss-avoidance learning processes. Specifically, while dopaminergic (DA) neurons, striatum and medial prefrontal cortex (mPFC) are generally involved in learning processes, sub-regions of the striatum (the dorsal part), as well as amygdala and anterior insula (AI) might be more intimately involved in punishment learning. Given these differences in neural mechanisms underlying reward and punishment learning, it is rational to hypothesize that the experimental treatment on learning performance is context-dependent. For instance, anxiety might selectively influence punishment learning as the amygdala is involved in threat detection and affective processes related to fear processing (Grupe & Nitschke, 2013; Trapp et al., 2018). Nevertheless, considering brain regions processing positive and negative affect might overlap (Bartra et al., 2013; Tom et al., 2007), these similarities in neural mechanisms can also lead to context-independent treatment effects. For example, the striatal-frontal circuit not only processes positive subjective value (Haber and Knutson, 2009; Engelmann et al., 2015), but also plays an important role in anxiety and metacognition (i.e., the variables of interest in the present thesis) (Grupe & Nitschke, 2013; Grupe, 2017; De Martino et al., 2013; Lebreton et al., 2015; Shapiro and Grafton, 2020). Although reward and punishment learning have been investigated at both the behavioral and neural level, the mechanism underlying the interaction between choice-irrelevant emotion (e.g., anxiety) and outcome valence and metacognition is still unclear.

1.3.3 Outcome valence and confidence judgments

Recent studies consistently found that confidence is altered by choice-relevant affect (Lebreton et al., 2018, 2019; Massoni, 2014; Sidi et al., 2018) and choice-irrelevant affect (Hoven et al., 2019; Koellinger & Treffers, 2015; Rouault et al., 2018) in both perceptual and value-based decisions. From the view of affective valence, people tend to overestimate the probability of positive event

(e.g., anticipated gains) and underestimate the probability of negative events (e.g., anticipated losses) (Giardini et al., 2008). This behavioral phenomenon is consistent with the current knowledge about confidence signals in the brain: confidence and valuation are automatically and generally processed in the same neural network (De Martino et al., 2013, 2017; Lebreton et al., 2015; Lopez-Persem et al., 2020; Shapiro & Grafton, 2020). Yet, to date the confidence-related brain regions were identified in the tasks using neutral or positive outcome. It remains unclear whether the same cognitive and neural systems are required for confidence processing when decisions are made under gain compared to loss contexts. Two pieces of evidence support that distinct systems might process confidence in different valence contexts. First, there is evidence for both overlapping but also distinct neural circuitry processing positive and negative subjective value (Bartra et al., 2013; Knutson et al., 2011; Oldham et al., 2018; Seymour et al., 2015; Tom et al., 2007). Given the overlap between confidence and valuation processes in vmPFC in the domain of gains (Lebreton et al., 2015; Lopez-Persem et al., 2020; Shapiro & Grafton, 2020), a similar overlap might be expected in the domain of losses. Second, previous fMRI results demonstrated that the neural mechanisms of confidence processing might be task-dependent (Sadeghi et al., 2017). In order to test these speculations, we optimized the learning task and ensured the robustness of valence effect on confidence judgment in the learning task first (**Chapter 3**). Afterward, we conducted a followed-up brain imaging study with the optimized learning task to investigate the neural mechanisms of confidence encoding in different affective contexts (**Chapter 4**).

1.4 Incidental Anxiety and Decision-Making

Another common source of affective information is the current emotional state at the time of choice, which may be carried over from choice-irrelevant events, such as mood, trait and induced choice-irrelevant affect. Although the role of emotions in decision-making has been extensively investigated, conclusions made by recent reviews about the role of anxiety in decision-making are still speculative (Grupe & Nitschke, 2013; Hartley & Phelps, 2012; Robinson, Vytal, et al., 2013; but see Grupe, 2017 for how anxiety involved in valuation). This section will briefly introduce anxiety and its relationship with stress and fear. Next, it will summarize the current understanding about the impact of anxiety on well-documented decision-making tasks including framing effects, loss aversion, risk aversion, and delay discounting. These findings demonstrate how anxiety also distorts (or not distorts) decision components (e.g., magnitude, probability and time) that are also involved in learning.

1.4.1 Anxiety

Around one third of the population suffers from an anxiety disorder at least once in their lifetime (Bandelow & Michaelis, 2015). Anxiety is usually associated with higher sensitivity to threat, intolerance of uncertainty, excessive worries and pessimistic thinking about future events (Cisler & Koster, 2010; Carleton et al., 2012; Holaway et al., 2006; Savitsky et al., 1998; Zenger et al., 2011; Dugas & Naomi, 2005). In addition to mental issues, anxiety is also associated with fear-related responses, such as muscle tension, increased sweating and heart rate. Moreover, anxiety is commonly accompanied by changes in the cognitive processes, some of which are likely important for decision-making. These anxiety-related changes in cognitive functions include attention, interpretation and working memory. For instance, anxious individuals tend to switch attention to threat stimuli and commonly interpret neutral events as threat (Grupe & Nitschke, 2013; Hartley & Phelps, 2012; Robinson, Vytal, et al., 2013). Moreover, impairments of working memory performance have been associated with both trait and induced anxiety across a number of studies, as found in a recent meta-analysis (Moran, 2016). Working memory is correlated with deep thinking, which is crucial for achieving long-term goals, such as learning and planning (Devetag & Warglien, 2003; Gill & Prowse, 2012). In addition, working memory also protects individuals from suffering from the impact of anxiety (Otto et al., 2013). The strong association between working memory, decisions and anxiety indicates that manipulations of working memory might further change choice patterns.

Research investigating anxiety and its treatments faces many challenges, primarily centered on the fact that the neural mechanisms of anxiety (i.e., subcortical regions) and defense mechanism of anxiety (i.e., physical and physiological responses to threat) recognized in animal studies are not enough to explain its interaction with unique cognitive functions that support complex decision-making seen in humans (Grillon et al., 2019; Grupe & Nitschke, 2013; LeDoux & Pine, 2016; but see Phelps & LeDoux, 2005). For example, Grupe and Nitschke (2013, 2017) proposed different functional connectivity between lateral cortical (e.g., insula), medial cortical (e.g., prefrontal cortex) and subcortical regions (e.g., amygdala) might be separately associated with anxiety-related alterations in decision-making in the context of uncertainty. In spite of the challenges that anxiety research faces, there is a growing body of literature developing innovative experimental designs

(Gillan et al., 2020; Grillon et al., 2019; Mkrtchian et al., 2017) and investigating the impact of anxiety in both clinically anxious and healthy participants. These efforts contribute to a growing field called computational psychiatry, which applies computational modelling to quantify psychiatric behavioral and neural changes. These quantitative changes are informative to better understand how our body (not just the brain) respond to challenging situations and to identify unusual response patterns that is likely to develop into a certain symptoms (see Huys et al., 2016, 2020; Montague et al., 2012 for more theoretical and practical applications).

1.4.2 Distinctions between Anxiety, Fear and Stress

Anxiety, fear and stress are considered distinct affective processes, even though they share several features. For instance, anxiety, fear and stress are triggered by threat events, they are associated with emotional arousal (i.e., accelerating heart rate, sweat) and evoke the same fear-related responses, such as fight, flight or freeze (Jelen et al., 2003; Roelofs, 2017; Roelofs et al., 2010). However, while anxiety, fear and stress are not mutually exclusive states and likely co-occur in everyday life emotional experiences, they also show distinct aspects that are not shared (Hartley & Phelps, 2012; Sylvers et al., 2011). Anxiety and fear can be distinguished through three aspects of the threat event that triggers affect, specifically the presence or absence of stimuli and hormone responses to threat. Firstly, fear is elicited by immediate threat events, and is therefore a reactive affective state, while future negative events are commonly associated with anxiety and stress, which are therefore anticipatory affective states. Secondly, fear and stress are induced by *predictable* threat (e.g., a deadline, social environment, spider), while anxiety is the response to *unpredictable* threat (Hartley & Phelps, 2012; LeDoux & Pine, 2016; Sylvers et al., 2011). The difference can be found at the neural level as well: the amygdala is suggested to process acute and predictable threat and the bed nucleus of the stria terminalis (BNST) is activates in response to unpredictable threat (Klumpers et al., 2017; LeDoux & Pine, 2016; Sylvers et al., 2011). Nevertheless, this idea was challenged by studies showing that activity in the amygdala is evoked by threat regardless of its predictability (Carlson et al., 2011; Hur et al., 2020). In reviewing these similarities and differences between anxiety, fear and stress, it is worth highlighting that it remains difficult to clearly distinguish between these emotional terms. One plausible reason is that emotion is a multidimensional construct and is unlikely to be differentiated by emotional terms (Lerner et al., 2015). This limitation also raises a question for self-reported emotion: can participants correctly identify similar but different emotions? To minimize this issue, one possibility is using valid emotion-induction techniques and ensuring the induced emotion is comparable to the symptoms of mental disorders defined by Research domain criteria (RDoC) or clinical criteria (e.g., DSM-5).

1.4.3 Stress and Anxiety Induction

Although stress and anxiety are not the same, , they are certainly related and findings on stress can inform those on anxiety, and vice versa (Berghorst et al., 2013; Bolton & Robinson, 2017; Clark et al., 2012; Gillan et al., 2020).One main reason is that acute stress might facilitate anxiety-like responses and therefore plays an crucial role in development of anxiety symptoms (Grillon et al., 2007). Moreover, anxiety and mental disorder caused by stress (e.g., post-traumatic stress disorder) are highly comorbid (Wisco et al., 2014). Accordingly, many literatures investigate the impact of acute stress on decision making, which might shed light on the mechanism of anxiety development.

In addition, there are two validated paradigms that have been used to induce acute stress: Cold-water Pressor Task (CPT; Porcelli et al., 2012) and the Trier Social Stress Task (TSST; Buckert et al., 2014; Kirschbaum et al., 1993). The former one requires participants to keep either right or left hand in the cold water (~3°C) before the task to induce immediate and physical pain. The later one requires subjects to prepare/give a talk in front of strangers for 5-15 minutes, which then induces social stress. These stress inductions usually take place before or in the middle of the tasks that assess the cognitive processes of interest. In these cases, experimenters are not able to track the emotional state during the task and the acute stress is less likely to develop into anxiety (Robinson, Vytal, et al., 2013).

To stick to the definition of anxiety, two validated anxiety induction techniques are commonly used to induce anxiety during the task. These two anxiety induction techniques are: Threat of shock (Engelmann et al., 2015; Grillon et al., 2004; Schmitz & Grillon, 2012) and low dose CO_2 inhalation (Bailey et al., 2011). The threat of shock paradigm (ToS) induces anxiety by delivering electric shocks at random time points throughout a task. Therefore, in this paradigm, the electric shock is unpredictable and uncontrollable during the task. Moreover, the ToS paradigm customizes the intensity of electric shock across participants, to ensure that each participant experiences emotional states of similar intensity toward the aversive event (i.e., electric shock). In contrast, the CO_2 paradigm simply requires participants to inhale hypercapnic gas (7.5% CO_2) during the task. These anxiety-induction paradigms have been widely used to investigate the mechanisms underlying the impact of anxiety on decision-making because they consistently trigger anxiety-like physiological (e.g., accelerating heart rate, sweat) and psychological responses (e.g., sensitive to the threat events) in healthy participants and further alter cognitive performance in ways comparable to the effect of pathological anxiety (Grillon et al., 2019; Hartley & Phelps, 2012; Robinson, Vytal, et al., 2013).

1.4.4 Impact of Anxiety on Value-Based Decision-Making

Decision-making requires multiple steps of information processing including representation and valuation. These steps feed into learning that updates option values (Rangel, 2009). Therefore, before investigating how anxiety is involved in learning process, I review the literature about the effects of anxiety on representation and valuation with four economic concepts: framing effects, loss aversion, risk aversion and delay discounting. Because research on the effects of anxiety on economic decision-making spans many fields from Social Psychology to Psychiatry and Medicine, only studies that fulfill the following criteria were included: (1) the same tasks were used across different anxiety modalities (i.e., trait, pathology and induced anxiety); (2) the measurement and tasks were well-defined to specifically study certain topic.

1.4.4.1 The effects of anxiety on framing

Given that anxiety is usually associated with biases toward negative perspectives and such feelings influence the representation of option (Gilovich et al., 2002; Slovic et al., 2007), anxious individuals would be predicted to show an enhanced framing effect as they are likely to be influenced by valence-framed options. This idea has been supported by findings from multiple studies quantifying the extent of the framing effect. The framing task we reviewed requires participants to select between risky and sure options after receiving an initial endowment at the beginning of each trial. The risky option consists of probabilistic gain and loss all of initial amount (e.g., 60% keep all; 40% loss all). The sure option is determined based on the expected value of risky option. For instance, if initial amount is $100 and the risky option consists of 60% keep all and 40% loss all, then the sure option will be "keep $40 from $100 for sure" in the gain-frame and "lose $60 from $100 for sure" in the loss-frame. The increase in framing-effect suggests that anxiety modulates the interpretation of contextual information (i.e., gain and loss), and selectively alters risk preferences rather than general risk avoidance. However, the result was not successfully replicated with the same gamble task by two recent studies, in which anxiety was induced via the threat of shock paradigm (Robinson et al., 2015) or measured as trait anxiety (Sip et al., 2016). Specifically, Sip et al. (2016) and Robison et al. (2015) found that anxiety generally reduces the selection of risky options. Besides the inconsistent effects of anxiety on option representations, anxiety might also/further target the evaluation of different components of choice options, including magnitude, valence probability and time. Next, the impact of anxiety on these components are separately reviewed.

Table 1.1 | Effect of anxiety on Framing effects

Author(s) and year	Threat Induction	Sample size	Anxiety effect	Effect size
Robinson et al., 2015	ToS (Within-subject)	N = 83 (M/F = 34/49)	=	$\eta^2 = 0.002$
Xu et al., 2013	Trait anxiety (Between-subject)	N = 20 (M/F = 9/11)	↑	r = 0.68
Sip et al., 2016	Trait anxiety (Between-subject)	N = 33 (OCD/HC = 18/ 16)	=	N/A
Gu et al., 2017	Trait anxiety (Between-subject)	N = 63 (M/F = 29/ 34)	↑	r = 0.32

White area indicates the anxiety is induced on healthy subjects. Yellow area indicates that trait/clinical measures of anxiety were employed.
=: no detectable impact. ↑: increase; ↓: decrease
OCD: Obsessive-compulsive disorder; **HC**: Healthy control; **GAD**: Generalized anxiety disorder

1.4.4.2 The effects of anxiety on loss aversion

Anxiety-induced changes in framing effects might be driven by an altered sensitivity to equivalent gains and losses. That is, anxiety generally increases negative affect triggered in response to outcomes and results in either lower sensitivity to gains or higher sensitivity to losses or both. This phenomenon can be depicted as an enlarged asymmetrical process of losses compared to gains (Kahneman & Tversky, 1979). To confirm this idea, I then review the literature about the effects of anxiety on loss aversion, which is commonly measured by fitting choices to random utility model (i.e., logistic regression) that yields a loss aversion parameter (Sokol-Hessner et al., 2009; Sokol-Hessner & Rutledge, 2019; Tom et al., 2007). Higher loss aversion indicates higher sensitivity to losses compared to equivalent gains. Surprisingly, only few studies used the paradigms mentioned above to directly examine the hypothesis that anxiety significantly changes loss aversion (Charpentier et al., 2016, 2017; Engelmann et al., 2015; Ernst et al., 2014) and, possibly even more surprising, all studies found little effect on behavioral loss aversion and conflicted with the hypothesis that anxiety would expand the asymmetrical processing of losses compared to gains. This conclusion has high external validity since these studies recruited participants from different ages (Adults: Charpentier et al., 2016, 2017; Ernst et al., 2014) and assessed not only trait and pathological, but also induced anxiety (trait anxiety: Charpentier et al., 2016; Pathological anxiety: Charpentier et al., 2017 and Sip et al., 2018; ToS-induced anxiety: Engelmann et al., 2015).

This surprising result may be attributed to the anxiety-specific neural circuits that process decision options. Specifically, Engelmann et al., 2015 found that subjective value processing is shifted from ventral medial prefrontal cortex (vmPFC) and ventral striatum (VS) to insula when risky decisions are made under induced anxiety. A dedicated neural circuit that processes subjective value under conditions of anxiety might not sufficiently influence other ongoing value computations and arrive at the same conclusions under certain conditions, for instance when the decision relies on the relatively simple comparison of gains vs. losses.

Table 1.2 | Effect of anxiety on Loss aversion

Author(s) and year	Threat Induction	Sample size	Range [Lambda]	Anxiety effect	Effect size
Engelmann et al., 2015	ToS (Within-subject)	N = 33	$-25 : $38 **[1.29]**	=	d =0.02
Ernst et al., 2014	Trait anxiety (Between-subject)	N = 66 (HC/GAD = 27/39)	$-20 : $40 **[1.39]**	=	d = 0.24
Charpentier et al., 2016	Trait anxiety (Between-subject)	N = 48 (M/F = 13/15)	$-10 : $18 **[1.56]**	=	r = -0.031
Charpentier et al., 2017	Patient study (Between-subject)	N = 48 (HC/GAD = 23/25)	$-10 : $18 **[2.0]**	=	d = 0.04
Sip et al., 2018	Trait anxiety (Between-subject)	N = 77 (OCD/HC = 43/34)	$-20 : $20 **[1.3]**	=	r =0.11

White area indicates the anxiety is induced on healthy subjects. Yellow area indicates that trait/clinical measures of anxiety were employed.
=: no detectable impact. ↑: increase; ↓: decrease
Lambda: Loss aversion parameter; **OCD**: Obsessive-compulsive disorder; **HC**: Healthy control;
GAD: Generalized anxiety disorder

1.4.4.3 The effects of anxiety on risk aversion

Subjective value estimates are based on a joint weighting of the magnitude, as well as the probability of a given outcome. As such, probability is another well-known choice component that is sensitive to the emotional state of the decision-maker (Hsee & Rottenstreich, 2004; Mukherjee, 2010). It may therefore also be impacted by anxiety. In order to specifically look into the effect of anxiety on attitude to known probability (so-called risk), the studies I review were selected with following criteria: tasks did not provide feedback about the chosen option until the end of the experiment, to avoid learning and history effects, and tasks used included one risky (i.e., probabilistic) and one sure option. The criteria were applied for two reasons. Firstly, feedback, or reward history, might influence decisions in different ways across participants (Schonberg et al., 2011). For instance, the Balloon Analogue Risk (BART) and Iowa Gambling Tasks (IGT) provide feedback at the end of each trial, which might change the strategy used by participants throughout the task and change the risk type investigated from a high level of ambiguity at the beginning of the task (probabilities are not known) to a lower level of ambiguity at the end of the task (probabilities are implicitly learned but cannot be expressed in numbers). Secondly, risk preference and probabilistic reasoning might covary with the outcome of previous performance. For instance, the empirical evidences showed that missing opportunity to gains and receiving substantial gains increases risk taking in the future (Büchel et al., 2011; Thaler & Johnson, 1990). These results reflect the essential feature of experience-based decision: subsequent choice is shaped by performance history. Meanwhile, these findings also point out potential confounding factors leading to difficulties of differentiating the "effect of anxiety on risk" from "effect of anxiety on learning or feedback processing". To "minimize" (not totally remove) this possibility, the potential solution is controlling the probability for at least one option (i.e., sure option) and avoiding the impact of feedback from previous decisions.

Risk attitudes are determined either by the proportion of sure options chosen or by utility functions which estimate a risk-aversion parameter. A higher value for either indices indicates less preference on the risky option than the sure option. The relationship between anxiety and risk attitude was extensively investigated. Using the same task paradigm (i.e., selecting between mixed-valence gamble and sure option), the majority of results showed that anxiety enhances risk aversion (Charpentier et al., 2017; Clark et al., 2012; Lempert et al., 2012). However, some studies reported little-to-no effect on risk attitude (Charpentier et al., 2017; Engelmann et al., 2015; Galván & Peris, 2014). These results indicated that the impact of anxiety on risk attitude is relatively stronger than the impact of anxiety on loss aversion[7].

[7] However, Mitte, 2006 using questionnaire found correlation between trait anxiety and risk attitude is modulated by subjective cost rather than subjective probability of negative events.

Table 1.3 | Effect of anxiety on Risk aversion

Author(s) and year	Threat Induction	Measures	Sample size	Anxiety effect	Effect size
Clark et al., 2012	ToS (Between-subject)	Risk avoidance	N = 65 (M/F = 35/30)	↑	$\eta^2 = 0.14$
Lempert et al., 2012	TSST	Risk avoidance	N = 113 (Male only)	↑	d = 0.30
Cohn et al., 2015	ToS (Within-subject)	Risk avoidance	N = 41 (N/A)	↑	N/A
Engelmann et al., 2015 Measurement 1	ToS (Within-subject)	Risk avoidance	N = 33 (Male only)	=	N/A
Engelmann et al., 2015 Measurement 2	ToS (Within-subject)	Risk aversion	N = 33 (Male only)	=	N/A
Galván & Peris, 2014	Trait anxiety (Between-subject)	Risk avoidance	N =31 (HC/GAD = 15/17)	=	N/A
Charpentier et al., 2017 Measurement 1	Patient study (Between-subject)	Risk avoidance	N = 48 (HC/GAD = 23/25)	=	d = 0.41
Charpentier et al., 2017 Measurement 2	Patient study (Between-subject)	Risk aversion	N = 48 (HC/GAD = 23/25)	↑	d = 0.72

White area indicates the anxiety is induced on healthy subjects. Yellow area indicates that trait/clinical measures of anxiety were employed. Risk avoidance refers to probability of taking sure option and risk aversion refers to the estimated parameter given choice pattern.
=: no detectable impact. ↑: increase; ↓: decrease
HC: Healthy control; **GAD**: Generalized anxiety disorder

1.4.4.4 The effect of anxiety on Delay discounting

Delay discounting refers to decisions that play out over time, such as the decision of whether to save money for a larger purchase later, or to spend it on a smaller item now. Usually, waiting pays off, as a larger value can be obtained at a later time. However, a longer delay between the decision and the outcome also increases the level of uncertainty. Delay discounting therefore measures the devaluation of future outcomes by asking subjects to make decision between immediate small and future larger outcomes. In line with anxiety enhancing the intolerance of uncertainty (i.e., treating unpredictable event as negative event; Behar et al., 2009; Buhr & Dugas, 2006; Dugas et al., 1995), it is reasonable to hypothesize that anxious individuals prefer immediate outcomes over future ones. Moreover, this hypothesis is further strengthened by the fact that anxious individuals have generally negative beliefs about the future. As **Table 1.4** shows, results concerning the impact of anxiety on delay discounting are mixed, with some studies showing decreases in patience (downward pointing error) and other no effect or increases.

Table 1.4 | Delay discounting

Author(s) and year	Threat Induction	Sample size	Anxiety effect	Effect size
Robinson et al., 2015	ToS (Within-subject)	N = 36 (M/F = 18/18)	=	$\eta^2 = 0.02$
Jenks & Lawyer, 2015	TSST (Between-subject)	N =113 (HAS/LSA = 50/63)	=	N/A
Rounds et al., 2007	Trait anxiety (Between-subject)	N = 88 (HAS/LSA = 50/63)	↓	N/A
Lempert et al., 2012	Trait anxiety (Between-subject)	N = 113 (Male only)	↓	d = 0.17
Engelmann et al., 2013	Patient study (Between-subject)	N =25 (HC/PTSD = 16/9)	↓	N/A
Lempert et al., 2015	Trait anxiety (Between-subject)	N = 45 (M/F = 17/28)	=	N/A
Jenks & Lawyer, 2015	Patient study (Between-subject)	N =113 (HAS/LSA = 50/63)	=	N/A
Steinglass et al., 2017	Patient study (Between-subject)	N =196 (HC/OCD/AN/SAD = 75/50/27/44)	↑ **(more patient)**	r = 0.24

White area indicates the anxiety is induced on healthy subjects. Yellow area indicates that trait/clinical measures of anxiety were employed.
=: no detectable impact. ↑: increase; ↓: decrease
OCD: Obsessive-compulsive disorder; **HC**: Healthy control; **GAD**: Generalized anxiety disorder

1.4.5 Brief Summary

Taken together, the current state of research suggests that the impact of anxiety on decision-making is not as strong as speculated by some researchers (Grupe, 2017; Hartley & Phelps, 2012; Robinson et al., 2015). A possible explanation for these differential results across studies is that these tasks and measurements reveal limited information about how anxiety is involved in the process of decision formation and its underlying mechanisms. This shortcoming has been strengthen by recent studies demonstrating a gap between risk attitude measured in laboratory settings and risk-taking in the field (Charness et al., 2020; Palminteri & Chevallier, 2018), indicating that the impact of emotion on choice patterns might be dynamic and context-dependent. By contrast, computational modelling captures inter- and intra-individual variability at particular time points throughout the experiment (Palminteri & Chevallier, 2018). Yet, the reliability and validity of computationally modelling has also been questioned. For instance, there is a main potential limitation when researchers implement prospect theory to interpret behavioral phenotypes (e.g., loss aversion, risk aversion, time discounting). The standard prospect theory models are static as they estimate parameters that do not allow for dynamic changes across time and contexts. In the other words, the previous performance is supposed to be independent of current decision, while this is not the case for most of real-life choices. This limitation leads to some artificial constraints on these parameters that will ignore the fact that performance history shapes decision (Büchel et al., 2011; Thaler & Johnson, 1990). To minimize this issue, we decided to implement a well-known reinforcement learning task, the two-armed bandit task, which allowed us to dynamically model the effect of emotion (in the present thesis, anxiety and outcome valence) on feedback processing and value updating as subjects learn to make advantageous decisions in a probabilistic context.

1.5 Outline of the Thesis

The main research question of the present thesis is whether and how learning process is influenced by affective contexts. To address this research question, the present thesis selectively investigates how incidental anxiety (i.e., choice-irrelevant affect) and outcome valence (i.e., choice-related affect) are involved in value- and confidence-updating. Before the empirical studies, the present chapter has summarized the current understanding about reinforcement learning, metacognitive function and affective contexts of interest in the decision-making field, which are relevant to our topics in chapters 2-4. In order to systematically investigate the impact of these two specific affective contexts on learning processes, I review the literature investigating effects of anxiety and outcome valence on value-based decisions and attempt identify shortcomings in experimental and theoretical approaches. Accordingly, I improve the experimental approach used for the experiments reported in chapters 2-4 and place results in the context of both model-free and model-based approaches to better understand learning processes in different affective contexts. The following chapters address our research questions from different standpoints:

Chapter 2 investigates how and whether incidental anxiety influences instrumental learning, while addressing and improving upon several issues with prior research identified in a focused literature review. We used a rich within-subject design, featuring both a learning and a transfer phase, and two affective framing manipulations: environment (anxiety vs safe) and outcome valence (gains vs losses). In two variants (N = 2x50) of this experimental paradigm, incidental anxiety was induced by delivering unpredictable, aversive and performance-independent electric shock during learning task. Moreover, the anxiety induction was assessed by both questionnaire and physiological responses (i.e., skin conductance responses; **Box 1**).

Chapter 3 focuses on the impact of choice-related affect (i.e., outcome valence: gain vs loss) on three common measurements: accuracy, confidence and reaction time (RT). While these three measurements are supposed to be highly correlated, it remains controversial whether valence-induced confidence and RT changes are dissociable. In order to address this issue, the goal of chapter 3 is to assess the presence of the valence-induced confidence bias in the absence of the RT bias. We conducted six variants of a learning task, attempted to disrupt the valence-induced motor bias effects by manipulating the mapping between decisions and actions and imposing constraints on response times (RTs).

Chapter 4 tests whether confidence formation in the brain is context-dependent or context-independent. We combined fMRI (**Box 2**) and a reinforcement learning task optimized in chapter 3, which dissociated gain and loss contexts and isolated motor responses from option evaluation. This combination enabled us to measure task-related brain activity and further identify the brain regions involved in confidence processing.

Box 1. Skin conductance response (SCR)
Unlike self-report questionnaire, SCR is recorded during the task and measures automatic nerve activity altered by emotional arousal regardless of awareness of emotion. SCR therefore serves as emotional marker to identify whether treatment successfully induces emotional responses. Like most of physiological data, the raw SCR is noisy and leads to difficulty in making conclusion based on raw data. Therefore, the repetition of the same event is needed to reduced noisy signal when data is average.

Box 2. functional magnetic resonance image (fMRI)
fMRI is the extension of MRI technique, which is common non-invasive brain measurement to investigate neural basis of decision-making. MRI is a three-dimension picture of brain structure regardless of neural activity. On the other hand, fMRI measures oxygen level(s) changed by event-induced neural activity and formed a four-dimension image (the fourth dimension is "time"). Neurons in the brain activate to process information. This procedure requires a blood flow to bring in red blood cell(s) with oxygen (i.e., oxygenated hemoglobin) and take away deoxygenated hemoglobin. Given different magnetic properties of oxygenated and deoxygenated hemoglobin, the proportion of oxygenated and deoxygenated hemoglobin causes blood-oxygenation level dependent response (i.e., BOLD signal) and is used to infer which brain region is involved in information processing. Considering each brain region might be responsible for multiple types of information processing, it is important to separate demands of cognitive functions via experimental design. If subjects are asked to make decision when options are displayed, it is difficult to dissociate motor-evoked or evaluation-evoked brain activation. Additionally, the timing of each event (e.g., displaying options) should be recorded to ensure which event (displaying options or motor response) elicits BOLD signal.

CHAPTER 2

The Effect of Anxiety on Reinforcement Learning[8]

[8] This chapter is based on Ting, C-C., Palminteri, S., Lebreton, M., Engelmann, J.B., The elusive effects of incidental anxiety on reinforcement-learning, PsyArXiv (2020a). *Under Revision.*

2.1 Introduction

The occurrence of negative events carries important information for an organism, enabling the adjustment of future behavior (Trapp et al., 2018). Yet, unpredictable negative events (e.g., Covid-19) can also cause prolonged anxiety and adversely impact otherwise well-adjusted behaviors, including decisions (Grupe, 2017; Hartley & Phelps, 2012; Robinson, Vytal, et al., 2013; Schmitz & Grillon, 2012). For example, anxiety is usually associated with higher risk-aversion (Charpentier et al., 2017; Clark et al., 2012; Cohn et al., 2015; Lempert et al., 2012), lower level of patient (Engelmann et al., 2013; Lempert et al., 2012; Rounds et al., 2007), as well as increased reliance on habituated behaviors (Browning et al., 2015; Raio et al., 2017; Schwabe & Wolf, 2009)[9]. Although habitual mechanism has its evolutionary role to make individual quickly response to the threat events, it might impair the flexibility of behavioral adjustment.

The ability to learn efficiently to seek rewards and to avoid punishments is one of the core features of adaptive behavior. Extensive evidence suggests that humans and animals learn by trial and error using algorithms akin to reinforcement learning, so as to repeat actions that maximize the occurrence of rewards and to suppress actions that lead to punishments (Rescorla & Wagner, 1972; Sutton & Barto, 1998). Given this pivotal role of reinforcement-learning in generating our behavior on the one hand, and the prevalence of anxiety in our daily lives on the other, a growing body of studies has investigated the impact of choice-irrelevant threat, the source of anxiety, and anxiety *per se* on learning in reward seeking and loss avoidance contexts (Abraham & Hermann, 2015; Berghorst et al., 2013; Browning et al., 2015; Cavanagh et al., 2011, 2019; DeVido et al., 2009; Glienke et al., 2015; Lighthall et al., 2013; Mather & Lighthall, 2012; Otto et al., 2013; Petzold et al., 2010; Robinson et al., 2013; Schwabe & Wolf, 2009; Treadway et al., 2017; see **Figure 2.1** and **Appendix A.1** for the mini-review).

To understand how previous literatures addressed this issue, we conducted a targeted literature review of key studies (**Figure 2.1**, N = 13, see **Appendix A.1** for inclusion/exclusion criteria) that investigated the impact of state (i.e., induced threat)[10] and trait anxiety on reinforcement learning. Both causal and correlational results from this literature review confirm the lack of consensus on the direction of the effects of anxiety on learning performance (**Figure 2.1**).

[9] See Chapter 1 for more examples about the impact of anxiety on decision-making.

[10] While anxiety mainly refers to the unpredictable aversive events, recent study also showed that no clear boundary in the brain to specify for predictable and unpredictable aversive events (Carlson et al., 2011; Hur et al., 2020). Moreover, many studies established hypothesis or interpreted results about the impact of anxiety on decision-making based on literatures investigating acute stress and anxiety (Berghorst et al., 2013; Bolton & Robinson, 2017; Clark et al., 2012; Gillan et al., 2020).

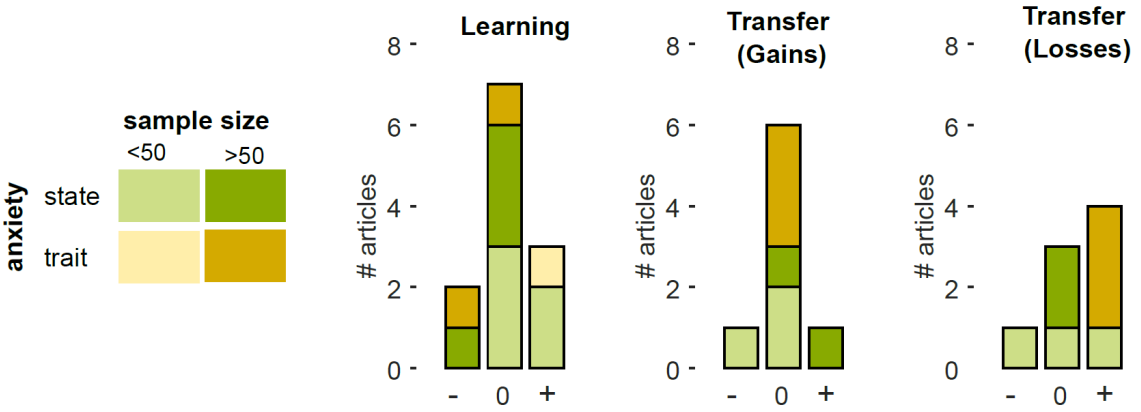


Figure 2.1. Overall effect of anxiety on learning and post-learning/transfer task. The stack bars summarized number of previous findings about the effect of anxiety on performance in learning task (middle-left panel), transfer tasks of approaching gains (middle-right panel) and transfer tasks of avoiding losses (right panel). In each pool of task, we separately reported numbers of articles showing decreases (-), no changes (0) or increases (+) of performance. The findings were categorized by both types of anxiety (green: state-anxiety; yellow: trait-anxiety) and sample size (brighter: N<50; darker: N>50). For instance, light green and light yellow represents effect of state-anxiety and effect of trait-anxiety with sample size N<50, respectively.

To resolve the seemingly contradictory effects of anxiety reported in prior research, we identified three main experimental dimensions which regularly differ between studies, and whose investigation could illuminate some of the discrepancies in the effects of anxiety on learning observed previously. These three dimensions, detailed in the following sections, cover 1) the methods of anxiety induction, including the nature of the stressor (i.e., aversive event), and the dynamics and intensity of the induced anxiety 2) the measures of learning performance, and 3) the manipulation of outcome valence.

Regarding the first factor, namely anxiety induction, a large variety of methods and protocols have been used in the literature. A significant proportion of studies investigating the impact of state anxiety on learning have used paradigms such as the Cold Pressor Test (CPT, Porcelli et al., 2012) and the Trier Social Stress Test (TSST, Jackson et al., 2006; Petzold et al., 2010), which suffer from significant drawbacks: these induction techniques operate before the learning task, making the actual emotion state less contingent with the task of interest and introducing uncertainty with respect to the dynamics of the emotion intensity and related endocrine reactions (Hermans et al., 2014; Robinson et al., 2013). Both self-reported anxiety and corticosterone (a glucocorticoid stress hormone) levels generally decrease over time after the initial stressor (Hermans et al., 2014; Jackson et al., 2006), suggesting that the peak of cortisol release, which is typically the time point at which the task is conducted, correlates with a state of relaxation that follows a stressful event, rather than a state of anxiety (Takahashi et al., 2005). Moreover, these stressors are not unpredictable, which is a critical factor of anxiety (Schmitz & Grillon, 2012). A significant portion of the studies

reviewed above might therefore not investigate the effects of anxiety, but instead complex post-stress recovery processes (Hermans et al., 2014). In the current experiments, we addressed this issue by using the well-established Threat of Shock (ToS) procedure to reliably induce anxiety during the learning task (Engelmann et al., 2015; Grillon, 2008; Schmitz & Grillon, 2012). ToS enables researchers to flexibly turn threat on and off (by contrasting periods during which electrical shocks are administered with periods of relative safety), which offers important advantages over other induction methods, including the ability to conduct experiments within participants and measuring in real time the causal effects of anxiety on decision-making.

Secondly, previous studies differ in what they refer to as *learning*. Two main experimental paradigms have been used to assess learning performance, which differ significantly in the aspect of learning they assess (see Palminteri & Pessiglione, 2017 for the comparison). More specifically, a first set of tasks (**Figure 2.1**, middle-left panel) primarily assesses the dynamic evolution of learning (*learning tasks*), while a second set of tasks (**Figure 2.1**, middle-right and right panel) mostly assesses post-learning preferences, much like extinction tests commonly employed in the animal learning literature (*transfer tasks*). Learning tasks directly assess the correct response rate during probabilistic instrumental-learning (see e.g., Pessiglione et al., 2006) and typically require participants to make repeated choices between fixed pairs of stimuli. Transfer tasks involve similar learning during an initial learning stage that provides feedback about the accuracy of participants' choices. However, learning performance is assessed after learning has already taken place in the form of an extinction test that involves novel pairings of the same stimuli and no longer includes feedback (Frank et al., 2004). Although these tasks seem very similar, those two ways of measuring learning performance have been shown to produce qualitatively different results, e.g. in the case of context-dependent learning (Klein et al., 2017; Palminteri et al., 2015). Accordingly, using different paradigms to capture the effects of anxiety on learning might not lead to comparable results across studies. We address this here by including both types of tasks in our experiments and separating these in our small-scale literature review, enabling us to assess the impact of anxiety on both learning and post-learning preferences.

Finally, despite the suggestions that the impact of anxiety could be valence-dependent, few studies have explicitly manipulated the valence of outcomes (gains vs losses) to contrast reward seeking and loss avoidance under conditions of anxiety (Berghorst et al., 2013; Cavanagh et al., 2011; Lighthall et al., 2013; Petzold et al., 2010). Instead, most studies have typically either limited their investigations and claims to one valence or re-framed low reward probabilities as an avoidance context (Schwabe and Wolf, 2009; Stevens et al., 2014). Neither of these approaches is actually suitable to investigate potential valence-specific effects of anxiety on learning (Palminteri & Pessiglione, 2017). We address this here by including both rewards and punishments in our experiments. This enables us to assess the differential impact of anxiety on reward seeking and punishment avoidance.

In the present study, we designed two experiments investigating the impact of anxiety on reinforcement learning to systematically address the shortcomings identified above. First, we employed Threat-of-Shock (ToS) to reliably and flexibly induce anxiety throughout the learning task (Cohn et al., 2015; Engelmann et al., 2015, 2019; Grillon et al., 2004; Schmitz & Grillon, 2012). In two different implementations of the task, we varied the dynamics and intensity of the anxiety induction, by applying Threat-of-Shock to relatively shorter blocks consisting of three trials or to

relatively longer periods consisting of the entire period of a learning session. In both experiments, shock intensity was calibrated for each individual. Second, we used a combination of tasks assessing both learning and transfer performance (Palminteri et al., 2015). Finally, we explicitly manipulated the valence of outcomes (gains and losses) to assess potential valence-specific effects of anxiety on learning (Pessiglione et al., 2006).

Regarding the analytical strategy, we first analyzed our data using standard linear mixed models that assess learning in different contexts on a trial-by-trial basis. Moreover, to more specifically assess how anxiety impacts on the underlying computations during learning and to parsimoniously make sense of this high-dimensional behavioral data, we used a recently developed computational modelling framework built around the concept of context-dependent learning (Palminteri et al., 2015; Palminteri, Lefebvre, et al., 2017). We aimed to identify the effects of anxiety on learning and its robustness across tasks and conditions: in other words, conditional on addressing what we identified as important caveats in previous studies (anxiety induction method based on ToS, explicit dissociation of learning and transfer performance, explicit gain and loss contexts), the anxiety effects should not be idiosyncratic to a specific experimental design, and should be comprehensively captured by computational modelling. Given reward and punishment learnings are suggested to require opposite motivations and different neural systems (Palminteri & Pessiglione, 2017; Pessiglione et al., 2006), the negative event might specifically influence learning when anticipated outcome is negative[11], we also specifically hypothesized that anxiety would impact context-dependent learning and/or valence-specific learning conditioned by context-dependent-learning. Despite our rigorous, comprehensive and high-powered experimental and analytical approaches, we found no clear, specific effect of anxiety on learning. In line with the lack of apparent consensus observed in the literature, our results seem to indicate that the effects of induced incidental anxiety on learning are at best elusive.

[11] However, some authors suggest that the induced threat (i.e., acute stress), the source of anxiety facilitates the sensitivity to positive feedback and consequently enhance gain learning (Lighthall et al., 2013; Mather & Lighthall, 2012).

2.2 Material and Methods

2.2.1 Participants

114 right-handed participants were recruited from the subject-pool of the Center for Research in Experimental Economics and Political Decision Making (CREED, www.creedexperiment.nl), and 100 participants were analyzed in the end (**Table 2.1**; Total: 56 males, aged 19-32, mean±SD = 23.27±3.08). We excluded four and ten participants from experiment 1 and experiment 2, respectively, either because of technical problems or average learning performance that was significantly lower than guessing level as identified via a binomial test assessing above chance performance (i.e., requiring a 50% performance at an alpha level of 0.01). All participants were prescreened via a questionnaire. Inclusion criteria consisted of (1) no history of psychiatric and neurologic disorders, (2) not taking medicine for anxiety or depression, (3) no implanted electric devices in the body (that electric shocks might interfere with), and (4) right-handedness. All participants gave their written informed consent before participation, after being given instructions about the task, the safety of electrical stimulation and their rights as participants. All procedures were executed in compliance with relevant institutional guidelines and were approved by the Economics and Business Ethics Committee (EBEC) at the University of Amsterdam.

2.2.2 Timeline of Procedure

We invited potential participants from the CREED subject-pool, and asked them to complete a battery of questionnaires at least one-day before the main task for an initial endowment of 10EU. When participants arrived at the lab, they were asked to thoroughly read the instructions and consent form and were allowed to ask questions to ensure understanding. We then orally explained the task if necessary. Participants' non-dominant hand (i.e., left hand) was then fitted with different electrodes meant to measure SCRs and deliver electric shocks. The successful setup was then followed by a calibration of shock intensity (see Anxiety induction), and a short training session (see behavioral task) while recording electrodermal activity. Subsequently the main task started and participants completed two (four) sessions in experiment 1 (experiment 2). Halfway through the learning experiments (i.e. before the second session for experiment 1, and before the third session for experiment 2), an additional calibration session was performed to control for (de)sensitization to the electrical stimulation. After the last learning session (the second session for experiment 1, the third and fourth session for experiment 2), participants completed the transfer task and an exit questionnaire. The total participant fee, including endowment amount and accumulated outcome from the learning task, was handed to participants in cash after completion of the exit questionnaire. The whole experiment took around 90 min, including instructions, electrodes setup time, exit questionnaire and payment (average amount earned in experiment1: mean±SD = 21.54 ±4.21; experiment2: mean±SD = 26.96 ±6.4).

2.2.3 Experimental Paradigm

All experimental paradigms were programmed and conducted with Matlab R2017b® (MathWorks) with the Cogent library (http://www.vislab.ucl.ac.uk/cogent.php).

2.2.3.1 Learning task

Participants performed a probabilistic instrumental learning task adapted from previous imaging, developmental and clinical studies (Palminteri et al., 2015; Palminteri et al., 2016; Salvador et al., 2017). They were instructed that the aim of the task was to maximize their payoff, by learning to choose the best cue out of cue pairs. They were explicitly told that seeking monetary rewards and avoiding monetary losses were equally important. Each learning session contained four novel, fixed pairs of cues, implementing a 2 (outcome valence: Gain vs Loss) x 2 (anxiety: Safe vs Threat) within-subject design. In other words, each pair of cues indicated a specific condition (**Figure 2.2B**; Safe/Gain, Safe/Loss, Threat/Gain, Threat/Loss). In the Gain conditions, possible outcomes were +0.5 or 0. Symmetrically, in the Loss conditions, possible outcomes were -0.5 and 0. The cue-outcome associations were determined by reciprocal but independent binomial probabilities, 75% or 25% (**Figure 2.2B**). Therefore, successful learning entailed choosing the cue associated with the higher probability of reward in the gain domain, and choosing the cue associated with the lower probability of loss in the loss domain.

Each block in experiment 1 (resp. of each session in experiment 2) started with a 1000ms reminder cue indicating the anxiety condition for the upcoming trials (i.e., "SHOCK" or "SAFE" associated with a frame of a particular color that was counterbalanced across participants – see **Figure 2.2A**) that was shown before the first trial. The reminder cues were followed by a fixation cross (1000-6000ms) and three trials. Each trial first featured a pair of cues (2500ms). During this cue display, participants indicated their decision by pressing the left or right arrow key to choose the left or right cue, respectively. The position of the options was counterbalanced. After 2500ms, an arrow appeared under the chosen cue (500ms). If participants did not respond in the allocated 2500ms, this phase was omitted and participants would get the relatively worse outcome in the feedback phase (i.e., -0.5 in the loss domain; 0 in the gain domain). In contrast, if participants successfully made decisions in time, the outcome associated with the chosen option was revealed (2500ms). Both trials and mini-blocks were separated by a jittered fixation cross (1000-5000ms).

In experiment 1, anxiety was induced through a ToS protocol used previous studies (Engelmann et al., 2015, 2019). More specifically, in order to maintain the emotional state for a prolonged period of time, we used a blocked presentation of the ToS conditions, such that three consecutive trials of the learning task were presented either under threat, or under safety (**Figure 2.2D**). Therefore, experiment 1 comprised two sessions, each including 96 trials (i.e. 32 blocks) and featuring a new, different set of eight cues. The ToS blocks were pseudorandomly interleaved to avoid repeating the same emotional treatment (Safe or Shock) more than two consecutive times. In experiment 2, we modified the experimental design and varied the ToS condition across separate sessions of 80 trials (with 20 repetitions of the four cue pairs; see **Figure 2.2D**) (Kim & Anderson, 2020). This was done in order to reduce the frequent switching of emotional states required by the relatively short blocks: the dynamics of emotion being notoriously slow (Williams et al., 2004), we wanted to exclude the possibility of spill-over effects of anxiety on Safe blocks. Consequently, only the valence of outcomes was manipulated within a session. Each session still featured four cue pairs, probabilistically associated with gains or losses. Experiment 2 comprised four learning sessions (two implementing the Safe condition and two implementing the threat condition, interleaved, with the order counterbalanced across participants).

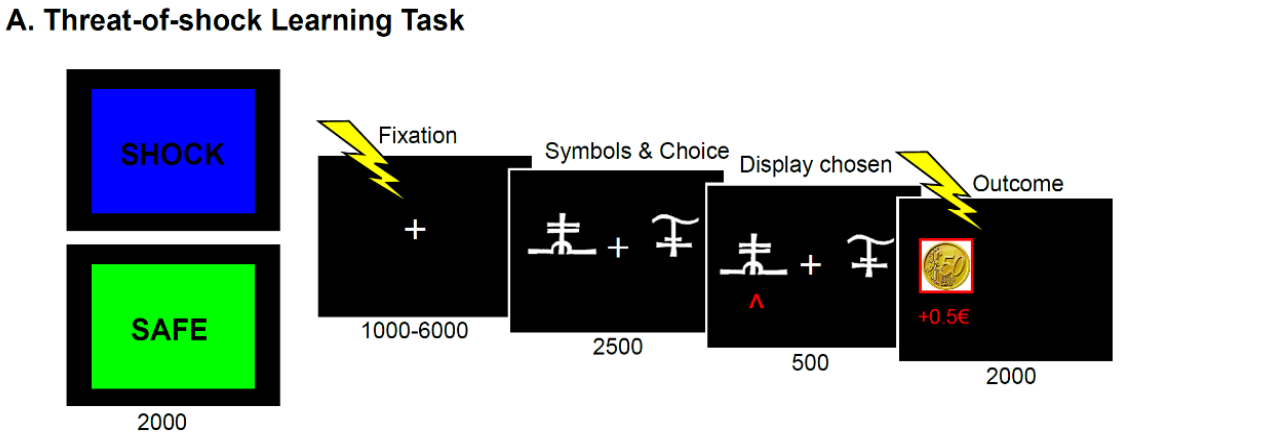

A. Threat-of-shock Learning Task
SHOCK
SAFE
2000
Fixation
Symbols & Choice
Display chosen
Outcome
1000-6000
2500
500
+0.5€
2000
B. Manipulations
Valence
Reward (+0.5€)
Punishment (-0.5€)
Threat of Shocks
Absent
Present
75% PGain 25%
25% PLoss 75%
C. Transfer Task
Fixation
Symbols & Choice
Display chosen
500-1500
3000
500
4 Learned pairs
24 New pairs
x 4 = 112 choices
D. Design: Session
Experiment 1: Full intermixed designs
2 Sessions
4 Contexts (A-D) / Session (1 / Condition)
24 Trials / Contexts
3 x consecutive trials same context
Experiment 2: Session manipulation
4 Sessions (2x Shock + 2x No-Shock)
4 Contexts (A-D)/ Session (2 / Condition)
24 Trials / Contexts
E. Design: Experiment
Experiment 1: Full intermixed designs
Learning Sess. 1
Learning Sess. 2
Transfer Sess. 1
Experiment 2: Session manipulation
Learning Sess. 1
Learning Sess. 2
Learning Sess. 3
Learning Sess. 4
Transfer Sess. 1
Transfer Sess. 2

Figure 2.2. Experimental paradigm. (A) Threat-of-Shock learning task. Schematic representation illustrating the learning task under safe and threat conditions for both experiment 1 and experiment 2. **(B)** Manipulations. In the two-by-two within-subject design, anxiety (i.e., absence and presence of threat of shock for safe and threat conditions, respectively) and outcome valence (i.e., gain and loss) were associated with specific pair of cues. Green and red represent gain and loss, respectively. Filled and unfilled rectangle represent safe and threat, respectively. Note that the cues were not framed with color in the experiment as figure (A) illustrates. **(C)** Transfer task. 24 novel cue pairings were formed by pairing each learned with all other cues in the transfer task. These new and original pairs were repeated four times, resulting in 112 choices. **(D)** Experimental design for each session in experiment 1 and experiment 2. *Top*: Experiment 1: Full intermixed design. The task contained cues from all four conditions. *Bottom*: Experiment 2: Session manipulation. Gain and loss from "either" safe or threat were associated with two pairs of cues, respectively. Therefore, four pairs of cues were used to represent two of four conditions. **(E)** Experimental design for both learning and transfer task. *Top*: Experiment 1: The transfer task took place right after second session and contained cues from all four conditions. *Bottom*: Experiment 2: The transfer task took place after the third and fourth sessions. The cues used in the transfer task were drawn from two of four conditions, depending on the emotional state of the previous session.

2.2.3.2 Transfer task

After completing the learning task, participants performed a transfer task (Palminteri et al., 2015). The task was built around the eight cues used in the last session(s) of the learning task. Participants were asked to choose between pairs of cues and indicate which cue they preferred. Yet, contrary to the learning sessions where cue pairs were fixed, all possible pairs were built from pairing each cue with the other 7 cues, leading to 28 combinations. Each pair was repeated four times, resulting in 112 trials (**Figure 2.2C**). Decisions were self-paced, and not followed by feedback about the decisions. Participants were not informed about the post-learning task until they had completed the learning task, so as to avoid explicit memorization strategies.

In experiment 1, the transfer task was conducted after the second learning session, which contained cues from both threat and safe conditions (**Figure 2.2E**, top). In experiment 2, the transfer task was conducted after the third *and* fourth session, to elicit choices between cues from both threat and safe conditions (**Figure 2.2E**, bottom).

2.2.3.4 Monetary compensation

In both experiments, all participants received a payment that included an initial endowment of 10 Euros for filling in the questionnaire before the experiment, a performance-based bonus based on all trials from the learning task (including gain and loss trials) and a final bonus (0.5€ for each question, totally four questions) for correctly filling out the exit questionnaire.

2.2.3.5 Anxiety induction

The incidental anxiety was induced by the presence of unpredictable and mildly painful electric shocks in the Threat conditions (Cohn et al., 2015; Engelmann et al., 2015, 2019; Grillon et al., 2004; Schmitz & Grillon, 2012). The shock stimulation was generated by a DS5 Isolated Bipolar Constant Current Stimulators (Digitimer Ltd.), and delivered through two electrodes. The electrodes were attached to the wrist of the non-dominant (left) hand throughout the experiment

where they were taped with Velcro. Calibrations for the intensity of electric stimulation took place three times in both experiment 1 and 2. In experiment 1, calibration occurred before the first and second learning sessions, and before the transfer task. In experiment 2, which consisted of four learning sessions, the calibration took place before the first and third learning sessions, and after the last transfer task.

The DS5 stimulator generated stable electric shocks with a fixed maximum input of 5V, maximum output of 25mA and a stable duration of 50ms across participants and studies. Only the intensity of shocks was individually customized for each learning session to match each subject's pain threshold. This was achieved using a staircase procedure asking participants to evaluate the painfulness of delivered electric shocks on a visual analog scaling ranging from 0 (not painful at all) to 10 (Extremely painful) (Engelmann et al., 2015, 2019; Story et al., 2013). The shock intensity was initialized at 10% of the maximum output (i.e., 2.5mA) and was then iteratively increased or decreased by 10% based on the following rules. If two consecutive ratings for the same intensity were less than 7, the intensity in the third trial would be increased by 10%. On the other hand, the intensity would be decreased when the rating was above 9. The procedure terminated as soon as the shock intensity was rated between 7 and 9 three times in a row, and we used this value in the subsequent learning sessions. During the calibration, the electric stimulations were self-triggered, i.e. participants could deliver the electric pulses by pressing the Enter key themselves. In order to avoid sensitization or desensitization, the intensity of electric stimulation was calibrated before and after each learning session in the experiment 1, and was calibrated before and after every two learning sessions in the experiment 2.

During the learning sessions, participants were not informed about the number of shocks and the precise time point of shock stimulation to maintain the unpredictability of electrical shocks. In Experiment 1, the number of shocks for each mini-block in the Threat Condition was randomly drawn from the pre-determined set [1 1 3 3 3 3 5 5] without replacement. The timing of the shock was randomized within the period of a threat block with the constraint that two consecutive shocks should be spaced more than 0.2 seconds apart. Participants were explicitly notified that the shocks were unpredictable, uncontrollable and independent of their performance. Similar procedures were applied to Experiment 2 with the exception that three electric shocks were delivered at random intervals within in each Threat session.

An important advantage of the ToS anxiety induction procedure is that anxiety states can be switched on and off during the task. To make sure that participants were subjected to this anxiety manipulation, mini-blocks (experiment 1) and sessions (experiment 2) started with a reminder (i.e. "SAFE" or "SHOCK") and a color frame (i.e. Green or Blue; independent of the valence factor, **Figure 2.2A**). The color frame was displayed until the end of the mini-block (experiment 1) or session (experiment 2). The assignment of color frames to Safe and Threat conditions was counterbalanced across participants.

2.2.3.6 Screening questionnaire

Data collected with the screening questionnaire was used to (1) pre-screen the subject by inclusion and exclusion criteria, and (2) assess a range of state and trait emotions. Exclusion criteria were examined first, and used to determine whether the subject qualified for the experiment (see **Section 2.2.1 Participants**). The screening questionnaire also included a basic demographic survey, Beck Depression Inventory (BDI: to index depression symptoms; Beck et al., 1988a), Beck Anxiety Inventory (BAI: to index clinical anxious symptoms; Beck et al., 1988b), State-Trait Anxiety Inventory (STAI: to index state- and trait-anxiety; Spielberger et al., 1983) and Positive and Negative affect schedule (PANAS: to index currently positive and negative affect; Watson et al., 1988). Finally, in order to avoid attentional biases from novel cues, the cues used in the experiment were displayed in the end of questionnaire for 60 seconds, so that participants have chance to explore them before the main task.

2.2.3.7 Exit Questionnaire

The exit questionnaire required participants to retrieve and report their emotional state on a 7-point scale and to explain the strategies they used in the task. For the self-reported emotions, participants were asked to separately rate how often they felt seven emotional-states (i.e., Anxiety, Fear, Happiness, Sadness, Anger, Surprise and Disgust) during the threat condition (from 0 (never) to 7 (every time)). In experiment 2, a few additional questions were added, where participants reported (1) the intensity of their emotional-state on the above emotions, and (2) their negative affect (from 0 to 7: positive to negative) and arousal level (from 0 to 7: arousal to calm) during both threat and safe conditions.

2.2.3.8 Skin Conductance Responses (SCRs)

SCR Acquisition

The SCRs were measured by Ag/AgCl electrodes filled with gel, and recorded via an amplifier and the software Vsrrp98 (version 7.29). After the instruction, two electrodes were attached on the ring and 3rd fingers of the left hand using medical tape. SCR data was collected at 1000Hz from the beginning to the end of the learning sessions (with the exception of 37 participants in experiment 1 for which a sampling rate of 500Hz was used). SCR data was synchronized with task events based on markers that indexed block/session onsets, trial onsets and feedback onsets.

SCR Analysis

Before statistical analyses, each participant's SCR data were preprocessed using the following steps: the data underwent (1) despiking by replacing outlier signals (defined as signals >= 3 times the standard deviation), (2) down-sampling to 10Hz and (3) normalization by z-scoring the data. The SCRs were analyzed as phasic responses relative to the trial onset (Bradley et al., 2000; Clark et al., 2012). Specifically, we extracted SCRs for the time window covering 2-4 seconds post trial onset and averaged the SCR response. From this we subtracted the trial-specific baseline, which was the mean SCR covering the 1-second period preceding trial onset (i.e., mean SCR during Inter-trial Interval). To avoid potential confounds caused by the delivery of the electric shock, the trials including shocks were not included in the analysis. Averages for each condition were then entered

into a two-way repeated measure ANOVA with outcome valence (gains vs. losses) and anxiety treatment (safe vs. threat) as within-subject factors. Note, that one (experiment 1) and five (experiment 2) participants were excluded from the SCR analysis, because of the low quality of the recorded SCR signal.

2.2.4 Behavioral Analysis

Learning task analysis

Correct choices from the learning task were extracted and served as a binary outcome variable (coded 1/ – respectively 0 – for a choice of the cue associated with the highest – respectively lowest – objective expected value). Averaged correct choice rates were computed per condition and per subject, and analyzed via a repeated-measures two-way ANOVA with (1) valence and (2) anxiety as factors, subject ID as a random effect, and a full interaction structure. To assess the effect of the differences in design between our two experiments on the different effects, we also added experiment number as between-subjects factor in the original ANOVA. Paired t-test were used to evaluate post-hoc comparisons between specific conditions.

As a complementary – and presumably more powerful – analysis, we also analyzed trial-by-trial data using a generalized linear mixed-effect (GLME) model. GLME models included independent variables accounting for the trial number (computed per condition, i.e. Experiment 1: trial = 0:1:23; Experiment 2: trial = 0:1:19), feedback valence (gain= 1; loss = -1), anxiety (threat = 1; safe =-1) and their 2- and 3-way interaction terms (valences* anxiety* trial). These variables were used in both the fixed-effects and the random-effects structure. The random effects structure accounted for the differences in experimental setups (coded 1 or 2 for Experiment 1 and Experiment 2, respectively) and inter-individual variations (subject's ID), which is nested within experiments. In Wilkinson-Rogers notation, this GLME writes as follows:

$$GLME1_{Learning}: Correct \sim (Intercept) + Valence * Anxiety * Trial + (1 + Valence * Anxiety * Trial \mid Experiment/Subject);$$

Given that the dependent variable was binary (the correctness of the choice) we used a logistic link function.

To directly assess the potential effects of the experimental designs on the manipulation effects, we also estimated the following GLMEs:

$$GLME2_{Learning}: Correct \sim (Intercept) + Valence + Anxiety * Experiment * Trial + (1 + Valence + Anxiety * Experiment * Trial \mid Subject);$$

Transfer task analysis

Similar to the learning task, the data from the transfer task was analyzed with both a repeated-measures ANOVA and GLME model. In order to impose similar data structures in Experiment 1 and in Experiment 2, we limited our analysis of Experiment 1 transfer pairs to trials where both cues were presented in the same Anxiety condition (i.e. both Safe or both Threat). Averaged choice

rates for each cue were computed, and analyzed using a three-way repeated-measures ANOVA with (1) option valence, (2) quality (option expected value: coded 1 if cue was the best of its pair during learning, 0 otherwise) and (3) anxiety manipulation as within-subject factors, subject's ID as a random effect, and a full interaction structure. Again, we added experiment number as between-subjects factor to the original ANOVA to account for the potential effects of design differences between our two experiments.

Because the preference relationship between intermediate values (i.e. Gain 25% -referred to as G25- and Loss 25% -referred to as L25) provide information about contextual learning (Palminteri et al., 2015), we ran additional analyses that focused on those cues. We submitted the choice rate of cues G25 and L25 to a two-way ANOVA with (1) option valence and (2) emotion manipulation. Afterward, we separately analyzed them for each comparison using a one-sample t-test.

Like for the learning task data, the transfer task data was further analyzed more comprehensively at the trial-by-trial level using a GLME approach. The model included independent variables accounting for differences between right and left cues, such as Diff_valence (difference in valence during learning) and Diff_quality (difference in the likelihood of avoiding a loss/attaining a gain during learning), and whether cues were learned in the context of threat (threat = 1; safe =0) and their interactions. These variables were entered into a logistic linear mixed model to predict binary choice based on the same structure of fixed- and random-effects, and also accounting for experiment (coded 1 or 2 for Experiment 1 and Experiment 2, respectively) and inter-individual variations (subject's ID), which is nested within experiment. In Wilkinson-Rogers notation, this GLME writes as follows:

$$GLME1_{Transfer}: ChooseRight \sim (Intercept) + Diff\_Valence * Diff\_Quality * Anxiety + (1 + Diff\_Valence * Diff\_Quality * Anxiety \mid Experiment/Subject);$$

Additional GLMEs added experiment number as fixed effect to assess its effect on experimental manipulation effects:

$$GLME2_{Transfer}: ChooseRight \sim (Intercept) + Diff\_Valence + Diff\_Quality * Anxiety* Experiment + (1 + Diff\_Valence + Diff\_Quality * Anxiety* Experiment \mid Subject);$$

All statistical analyses were performed using Matlab R2015a. GLME models were estimated using the function fitglme.

2.2.5 Computational Modelling

Step 1: identifying the best computational architecture

In a first modelling stage, we aimed to identify the general algorithm governing learning, regardless of the anxiety condition. Following a previous approach (Rescorla & Wagner, 1972), we first built a nested model-space (Model Space 1), including six increasingly complex RL models (see **Figure 2.6 A-B** for the illustrations). The six models are referred to as ABS, REL, REL_w, ABS_a and REL_a, and $REL_{a,w}$, where REL and ABS respectively referred to ABSOLUTE and RELATIVE, 'a' to

asymmetric, and 'w' to weighted counterfactual outcome. The ABSOLUTE and RELATIVE models were introduced in (Palminteri et al., 2015).

In the ABSOLUTE model, at each trial t, the chosen option value (c) of the current context *s* is updated with the Rescorla-Wagner rule (Rescorla and Wagner, 1972):

$$Q_{t+1}(s,c) = Q_t(s,c) + \boldsymbol{\alpha} \times \delta_t \qquad \text{Equation (1)}$$

Where $\boldsymbol{\alpha}$ is the learning rate for the chosen option and δ_t is the prediction error term calculated as follows:

$$\delta_t = R_t(s) - Q_t(s,c) \qquad \text{Equation (2)}$$

In the RELATIVE model, a choice context value ($V(s)$) is also learned and used as the reference point to which an outcome should be compared before updating option values.
Context value is also learned via a delta rule:

$$V_{t+1}(s) = V_t(s) + \boldsymbol{\alpha_V}\, \delta_{V,t} \qquad \text{Equation (3)}$$

Where $\boldsymbol{\alpha_V}$ is the context value learning rate and δ_V is a prediction error-term calculated as follows:

$$\delta_{V,t} = (R_t(s) + \neg R_t(s))/2 - V_t(s), \qquad \text{Equation (4)}$$

$\neg R_t(s)$ indexes the outcome not received available in context *s*, and is computed as follows:

$$\neg R_t(s) = \begin{cases} 0 \; if \; R_t(s) = -0.5 \text{ or } 0.5 \\ 0 \; if \; R_t(s) = 0 \text{ and } V_t(s) = 0 \\ 0.5 \; if \; R_t(s) = 0 \text{ and } V_t(s) > 0 \\ -0.5 \; if \; R_t(s) = 0 \text{ and } V_t(s) < 0 \end{cases} \qquad \text{Equation (5)}$$

Therefore $\neg R_t$ captures the fact that participants infer that the non-selected cue is associated with the complementary outcome to the one they actually received. The formulation of $\neg R_t$ depends on the context value $V_t(s)$ because context values have to be disambiguated (i.e. gain or loss context) before participants can infer the outcome that is complementary to 0. Note that this specification slightly differs from the original model proposed in Palminteri et al., (2015), which writes:

$$\delta_{V,t} = (R_t(s) + Q_t(s,u))/2 - V_t(s), \qquad \text{Equation (6)}$$

where $Q_t(s,u)$ is the Q-value of the unchosen option.

The proposed modifications to the RELATIVE model (Equations 4-5) are meant to account for the significant context dependency observed in our data, evidenced in the Transfer task data (See Results - Model-based analysis indicates that learning is asymmetric and context-dependent). In addition, this formulation provided a better fit of the data than the original one in a formal model-comparison.

In the asymmetric models (ABS_a, REL_a, $REL_{a,w}$), we additionally introduced different learning rates after positive versus negative prediction errors. This follows from several studies showing that individuals tend to give more weights to positive, confirmatory feedback than to negative, disconfirmatory feedback (Lefebvre et al., 2017, Palminteri et al., 2017).
In those models, Equation (1) therefore becomes

$$\begin{cases} Q_{t+1}(s,c) = Q_t(s,c) + \boldsymbol{\alpha}^{+} \times \delta_t \; if \delta_t > 0 \\ Q_{t+1}(s,c) = Q_t(s,c) + \boldsymbol{\alpha}^{-} \times \delta_t \; if \delta_t < 0 \end{cases}$$

Equation (7)

In the weighted models (REL_w, $REL_{a,w}$), the inference on the forgone outcome $\neg R_t(s)$ was modulated by a weight $\mathbf{w}$ as follows:

$$\delta_{V,t} = (R_t(s) + \mathbf{w}(\neg R_t(s)))/2 - V_t(s), \qquad \text{Equation (8)}$$

In all models, the probability of choosing option *A* over *B* was derived from a softmax function with temperature parameter β:

$$P(choice = A) = (1 + \exp(\boldsymbol{\beta}(v(A) - v(B))))^{-1} \qquad \text{Equation (9)}$$

In the learning task at trial t, we have, for an option i of a context s: $v(i) = Q_t(s,i)$
In the transfer task, we have, $v(i) = Q_{end}(s,i)$, where $Q_{end}(s,i)$ indicate the Q-values of option i at the end of the learning session.

Step 2: Modelling the effects of anxiety

After having identified the general algorithm governing learning, we next investigated if and how incidental anxiety – as induced by threat of shocks – affects specific sub-processes of learning. We therefore defined a second model space (Model Space 2) by systematically allowing each parameter (the temperature parameter, each of the three learning-rates, and the weighting parameter) of the winning model (i.e., model $REL_{a,w}$) to differ between the safe and the threat condition. This produced a 5-models model-space, to which was added a base model where all parameters were identical between the safe and the threat condition (**Figure 2.8A**).

Initialization
Option (*Qs*) and context (*Vs*) values were initialized at 0 in each condition.

Parameter optimization
For each model M, and regardless of the criterion used for model comparison (see equations below), the parameters θ_M were optimized by minimizing the negative logarithm of the posterior probability (*LPP*) over the free parameters:

$$\text{LPP} = -\log\big(P(\theta_M|D,M)\big) \propto -\log\big(P(D|M,\theta_M)\big) - \log(P(\theta_M|M)) \qquad \text{Equation (10)}$$

Here, $P(D|M,\theta_M)$ is the likelihood of the data (i.e. the observed choice) given the considered model M and parameter values θ_M, and $P(\theta_M|M)$ is the prior probability of the parameters. Following Daw et al., (2011), the prior probability distributions were defined as a gamma distribution with two parameters of 1.2 and 5 (which is written as gampdf(β,1.2,5) in Matlab) for the choice temperature (β), and as beta distributions with two parameters of 1.1 (which is written as betapdf(α,1.1,1.1) and betapdf(w,1.1,1.1) in Matlab) for learning rates (α) and weight (w).

This procedure was conducted using Matlab's fmincon function with different initialized starting points of the parameter space (i.e., 0<β<Infinite, 0<α<1) (Palminteri et al., 2015). Note that both the learning and transfer task data were used for the parameter optimization.

Model comparison criteria

We computed three model comparison criteria, which measure the ability of each model to explain the experimental data, by trading-off their goodness-of-fit and complexity: the Bayesian Information Criteria (*BIC*), the Akaike's Information Criteria (*AIC*) and the Laplace approximation to the model evidence (*LAME*).

Defining $\hat{\theta}_M$ the model parameters identified in the optimization procedure, *df* the number of model parameters, and *n* the number of data-points (i.e. trials), AIC, BIC and LAME were computed as follows:

$$\mathrm{BIC} = \log\left(P\left(D\middle|M,\hat{\theta}_M\right)\right) - \frac{df}{2}\log(n)$$

$$\mathrm{AIC} = \log\left(P\left(D\middle|M,\hat{\theta}_M\right)\right) - df$$

$$\mathrm{LAME} = \log\left(P\left(D\middle|M,\hat{\theta}_M\right)\right) + \log\left(P\left(\hat{\theta}_M\middle|M\right)\right) + \frac{df}{2}\log(2\pi) - \frac{1}{2}\log|H|$$

Where $|H|$ is the determinant of the Hessian.

These three criteria were compared in their ability to correctly identify model simulations (see Model identifiability and parameter recovery section below). Because LAME gave the most satisfactory results, only model comparisons using this criterion are reported in the main text.

Bayesian model comparison

To identify the model most likely to have generated a certain data set, AIC, BIC and LAME were computed at the individual level for each model in the respective model-space, and fed to random-effects Bayesian Model Comparison using the mbb-vb-toolbox (http://mbb-team.github.io/VBA-toolbox/; Daunizeau, Adam, and Rigoux, 2014). This procedure estimates the expected frequencies (denoted PP) and the exceedance probability (denoted XP) for each model within a set of models, given the data gathered from all participants. XP quantifies the belief that the model is more likely than all the other models of the model-space. An XP >95% for one model within a set is therefore typically considered as significant evidence in favor of this model being the most likely. Expected frequency (PP), on the other hand, quantifies the posterior probability, i.e., the probability that the model generated the data for any randomly selected subject.

Model identifiability and parameter recovery

In order to assess the reliability of our modelling approach, we performed model identifiability and parameter recovery simulations (see Correa et al., 2018 for a similar approach). Choices from synthetic participants were generated for each task and each model by running our computational models with model parameters sampled in their prior distribution: softmax temperature β were drawn from gamma distribution (i.e., random('Gamma',1.2,5) in Matlab) and learning rates and weights were drawn from beta distributions (i.e., random('beta,1.1,1.1) in Matlab), as outlined above. Option values Qs and context values Vs were initialized from 0 for four conditions. For each model, we ran 10 simulations including 50 synthetic subjects (N=500).

Model identifiability was assessed by running the Bayesian Model Comparison on the synthetic data. Results are pictured as confusion matrices, where perfect recovery would result in matrices with diagonal elements equal to 1, and off-diagonal elements close to 0. Parameter recovery was assessed by evaluating the correspondence between the parameters used in the simulation, and the parameters recovered by the parameter optimization procedure (**Figure 2.6-2.8**). We used two main assessment criteria: first, we performed a linear regression analysis between the parameters used for simulations and the estimated parameters, using data from all simulations (n = 500). In this case, perfect recovery would result in intercepts close to 0 and slopes close to 1, and would be pictured as a 500 dots scatter plot aligned on the identity line. Then, we performed correlation analyses between the parameters used for simulations and the estimated parameters on individual simulations (each with n = 50 synthetic data). Correlation coefficients were averaged over the 10 simulations, and displayed as correlation matrices. In this case, perfect recovery would result in matrices with diagonal elements equal to 1, and off-diagonal elements close to 0.

2.3 Results

2.3.1 Manipulation Checks: successful induction of anxiety

In order to ensure that our anxiety manipulation was successful, we inspected self-reported emotion and physiological responses. In both experiments, self-reported anxiety during the threat condition was significantly higher than other (negative and high arousal) emotions, including sadness (Exp.1: $t_{44} = 6.78$, $p < .001$, $d = 1.01$; Exp.2: $t_{49} = 6.17$, $p < .001$, $d = 0.87$) and anger (Exp.1: $t_{44} = 3.56$, $p < .001$, $d = 0.53$; Exp.2: $t_{49} = 4.98$, $p < .001$, $d = 0.70$), but was similar to fear (Exp.1: $t_{44}=1.57$, $p =.1235$, $d = 0.23$; Exp.2: $t_{44} = 0.67$, $p = .5024$, $d = 0.09$). While self-reported anxiety levels were significantly greater than 0 in both experiments (**Table A.2.1**), anxiety levels were significantly higher in experiment 1 compared to experiment 2 ($t_{93}= 4.28$, $p < .001$, $d = 0.87$). This indicates that blocking ToS, as done in experiment 1, induced greater levels of anxiety compared to the more prolonged ToS presentation in experiment 2.

The self-report results are paralleled by the SCR results, which showed significantly higher phasic responses during anxious compared to safe trials (**Figure 2.3**). An ANOVA showed a main effect of *anxiety* (Exp.1: $F_{1,47} = 130.35$, $p < .0001$, $\eta_p^2 = 0.69$; Exp.2: $F_{1,44} = 4.28$, $p = .0444$, $\eta_p^2 = 0.08$). This effect was not modulated by outcome valence (Exp.1: $F_{1,47} = 0.08$, $p = .7719$, $\eta_p^2 = 0.02$; Exp.2: $F_{1,44} = 0.00$, $p = .9901$, $\eta_p^2 = 0.00$). While SCR levels were significantly greater during threat compared to safe trials in both experiments (Exp.1: $t_{48} = 10.44$, $p < .0001$, $d = 1.49$; Exp.2: $t_{44} = 2.06$, $p = .0444$, $d = 0.30$), arousal levels during threat (vs. safe) conditions were also significantly greater on average in experiment 1 compared to experiment 2 (Exp.1> Exp.2: $t_{92} = 6.94$, $p < .0001$, $d = 0.35$). Jointly, results from SCR and self-report indicate that ToS successfully induced anxiety in both experiments.

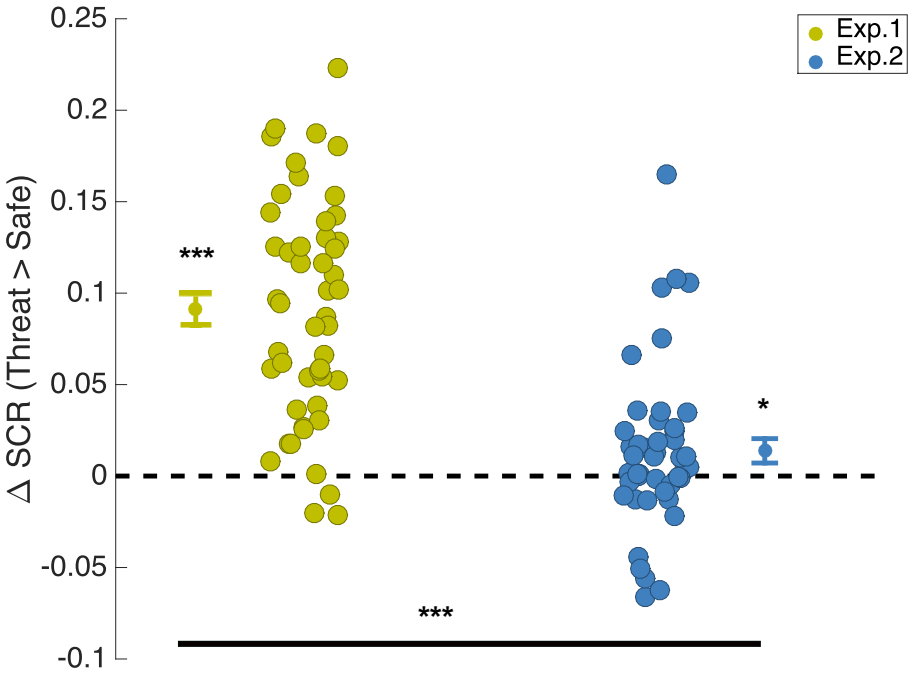


Figure 2.3. Electrophysiological results. Skin conductance responses (SCR) were significantly higher in the threat condition compared to the safe condition (Exp.1: t_{48} =10.44, P < .0001, d = 1.49; Exp.2: t_{44} =2.06, p = .0444, d = 0.30). Moreover, this effect was more significant in Experiment 1 than Experiment 2 (Exp.1> Exp.2: t_{92}= 6.9499, p < .0001, d = 0.35). The dots represented the mean of ΔSCR and the error bars represented the standard error of the mean.
~ .05< p <.10; * .01< p <0.05; ** .001< p < .01; *** p< .001

2.3.2 Model-Free Analysis: No effects of valence and anxiety on learning performance

To investigate the overall effects of our experimental manipulations (valence: gain vs. loss; anxiety: safe vs. threat) on learning performance, we first analyzed the probability of correct responses averaged per condition, using a two-way repeated-measures ANOVA. This analysis revealed no significant main effects of - nor significant interactions between - our experimental factors on learning performance (Exp.1: Ps' > 0.12; Exp.2: *ps* > 0.14; see **Figure 2.4A** and **Table A.2.2-A.2.3**). While the two variants of the ToS procedure had a marginally different impact on average learning performance (Exp.1 = 74.12%, Exp.2 = 71.60%; $F_{1,99} = 2.66$, $p = .1057$, $\eta_p^2 = 0.02$), they did not induce significantly different effects of anxiety on learning ($F_{1,99} = 0.008$, $p = .9271$, $\eta_p^2 = 0.00$). As subsequent analyses focus on identifying the effect of anxiety, we therefore combined the data from the two experiments (but we additionally continue to report individual experiment results for all main analyses).

Combining the two experiments in a single ANOVA, we replicate the absence of significant main and interaction effects on learning performance (All: *ps* > 0.20; **Figure 2.4A** and **Table A.2.3**), as reported above. Although the lack of valence effects replicates previous findings (Fontanesi et al., 2019; Lebreton et al., 2019; Palminteri et al., 2015), the absence of significant effects of anxiety might seem surprising at first glance, as it contradicts several previous studies suggesting anxiety affects learning *per se* (DeVido et al., 2009; Glienke et al., 2015; Treadway et al., 2017). However, our findings agree with our literature review of 13 studies on the effects of threat states and trait anxiety on instrumental learning, which also failed to identify a consensual, robust effect of anxiety on learning (**Figure 2.1** and **Table A.1**).

Because the absence of effects of anxiety contradicts our a priori hypothesis, we next turned to a more flexible statistical analysis framework using a generalized linear mixed-effect (GLME) model. This approach allows us to inspect our data trial-by-trial to capture learning effects and may be more powerful than ANOVAs in the presence of unbalanced or missing data (Matuschek et al., 2017; Pinheiro & Bates, 2000). Although the GLME revealed a main effect of trial on performance, capturing the dynamics of learning (**Figure 2.4B** and **Table A.2.4**), no other significant main effects and/or interaction with the experimental manipulations were detected. Confirming the ANOVAs results, this indicates that threat of shock might have a limited impact on learning processes. Moreover, we also did not find an effect of experimental designs on learning ($\beta_{exp} = -0.00 \pm 0.03$, $t_{25276} = -0.14$, $p = .8834$; $\beta_{exp*anxiety} = 0.03 \pm 0.04$, $t_{25276} = 0.82$, $p = .4082$; see **Appendix A.3**).

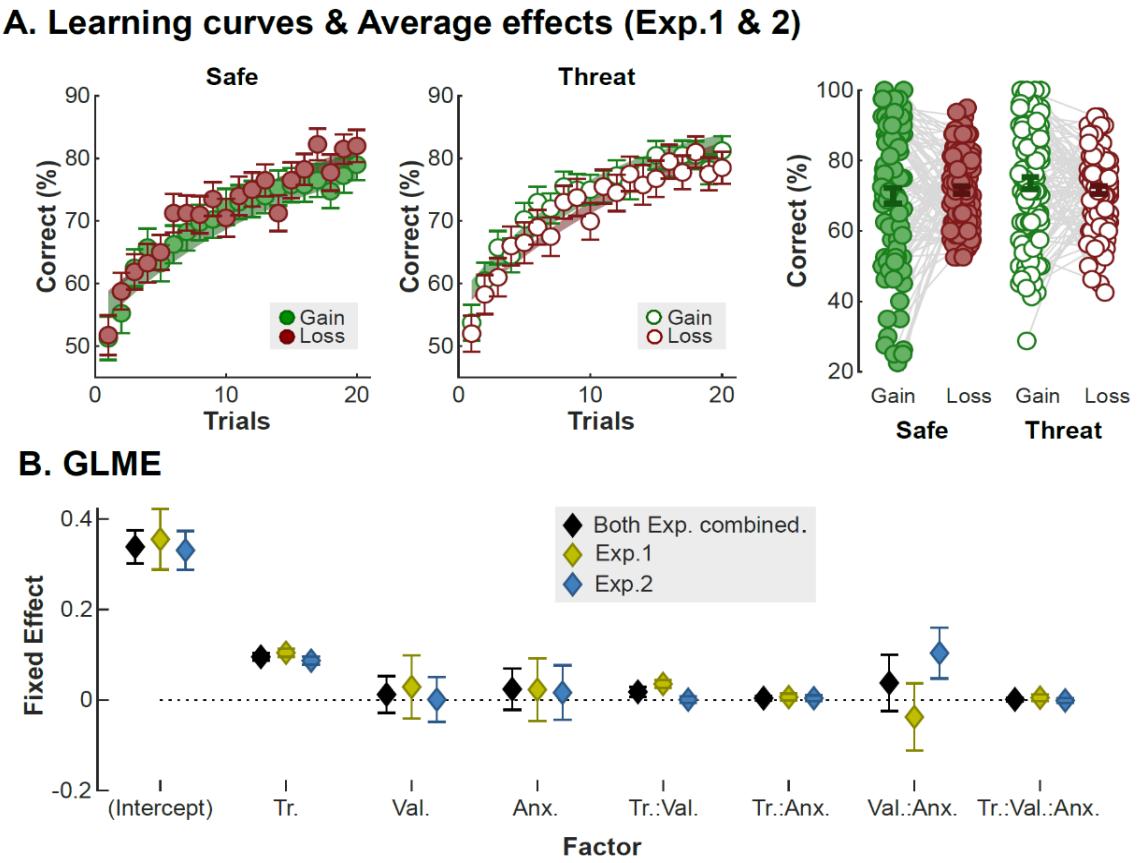


Figure 2.4. Learning performance and results from generalized linear mixed-effect model. **(A)** Left and middle panels: learning curves representing the fraction of correct choices in the safe (left; filled dots) and the threat (middle; unfilled dots) conditions. Error bars indicate the standard error of the mean (SEM), and shaded areas represent the mean ± SEM of GLME predictions. Right panels: average correct rate across four conditions. Each gray line indicates individual's choice patterns across the conditions. **(B)** Generalized linear mixed-effect model (i.e., $GLME1_{Learning}$) with choice accuracy as dependent variable. The y-axis represents the estimated standardized coefficient (t-value), and the x-axis represents each factor in the GLME model. Dot colors indicate results from different dataset. **Tr.**: Trial; **Val.**: outcome valence; **Anx.**: Anxiety manipulation

2.3.3 Model-Free Analysis: non-specific effects of anxiety on transfer task performance reflecting learned values

Participants' choices in the transfer task provide additional information about the value of the cues that they have learned throughout the learning task. We first computed the preference for each cue as the probability of choosing the cue over all other cues (**Figure 2.5**; and see Palminteri et al., 2015 for a similar approach). Importantly, the three-way repeated-measures ANOVA (**Table A.2.5** and **Section 2.2.4: Methods – Behavioral analysis**) identified an interaction between cue quality and anxiety (interaction of anxiety and quality; $F_{1,99} = 6.37$, $p = .0131$, $\eta_p^2 = 0.06$). Post-hoc tests were performed to characterize the interaction between anxiety and quality. Results indicate that the interaction is driven by participants choosing the better symbols (G75 and L25) over the worse symbols (G25 and L75) significantly more often in the threat compared to the safe condition (Better-Worse in threat > Better-Worse in safe: $t_{99} = 2.41$, $p = .0165$, $d = 0.17$; **Table A.2.8**). These results indicate that anxiety boosts participants' preference for higher quality cues.

Additional results from this ANOVA revealed that a cue is more likely to be preferred if it was associated with gains compared to losses (main effect of option valence; $F_{1,99} = 212.85$, $p = 2.07 \times 10^{-26}$, $\eta_p^2 = 0.68$), regardless of the anxiety condition ($F_{1,99} = 0.31$, $p = .5756$, $\eta_p^2 = 0.00$). A cue is also more likely to be preferred if it was the best cue (G75 and L25) of the pair during the learning task (main effects of Option quality; $F_{1,99} = 241.60$, $p = 2.59 \times 10^{-28}$, $\eta_p^2 = 0.70$). However, there were no valence-dependent effects of anxiety on preferences in the transfer task ($F_{1,99} = 0.55$, $p = .4586$, $\eta_p^2 = 0.00$), as well as no main and interaction effects of anxiety with the factor experiment ($ps > 0.19$). The latter result suggests that the two anxiety induction methods did not differentially impact preference in the transfer task. Taken together, anxiety during learning improved recognition of cue quality in the transfer task independent of valence.

Those results were confirmed in a more comprehensive GLME approach (see **Section 2.2.4 Material and Methods – Behavioral analysis**), which modelled transfer task choices between two cues as a function of (1) the difference between the cues' valence (gains vs. losses), (2) the difference between the cues' quality (better option in learned pair vs. worse option in learned pair), and (3) whether cues were learned in the anxiety condition (safe vs shock). The GLME model also accounted for differences in experimental designs (experiment 1 vs. experiment 2) and subject ID, which was nested within each experiment. The results (**Figure 2.5B** and **Table A.2.6**) showed that participants' decisions were influenced by valence ($\Gamma_{13592} = 8.90$; $p = 5.83 \times 10^{-19}$), quality ($\Gamma_{13592} = 6.53$; $p = 6.73 \times 10^{-11}$) and the interaction between anxiety and quality ($\Gamma_{13592} = 2.44$; $p = .0144$). Comparable to the ANOVA results above, we did not find a main effect of experiment ($\beta_{exp} = -0.01 \pm 0.08$, $t_{13591} = -0.13$, $p = .8940$) nor an interaction between experiment and anxiety ($\beta_{exp*anxiety} = -0.03 \pm 0.12$, $t_{13591} = -0.24$, $p = .8041$; see **Appendix A.3**), indicating that the experimental design had little effect on the impact of anxiety on post-learning performance.

Post-hoc analyses showed that cue discrimination (G75+L25 > G25+L75) was significantly improved in the threat compared to the safe condition ($t_{99} = 2.41$, $p = .0165$, $d = 0.17$). Note that this effect was not significant in experiment 1 ($t_{49} = 1.27$, $p = .2058$, $d = 0.12$), but the direction was the same as observed in experiment 2 ($t_{49} = 2.42$, $p = .0171$, $d = 0.24$). The results indicate that anxiety might enhance participants' ability to identify the higher quality symbol some time

after learning, even though it does not affect average learning performance at the learning stage (**Figure 2.4**).

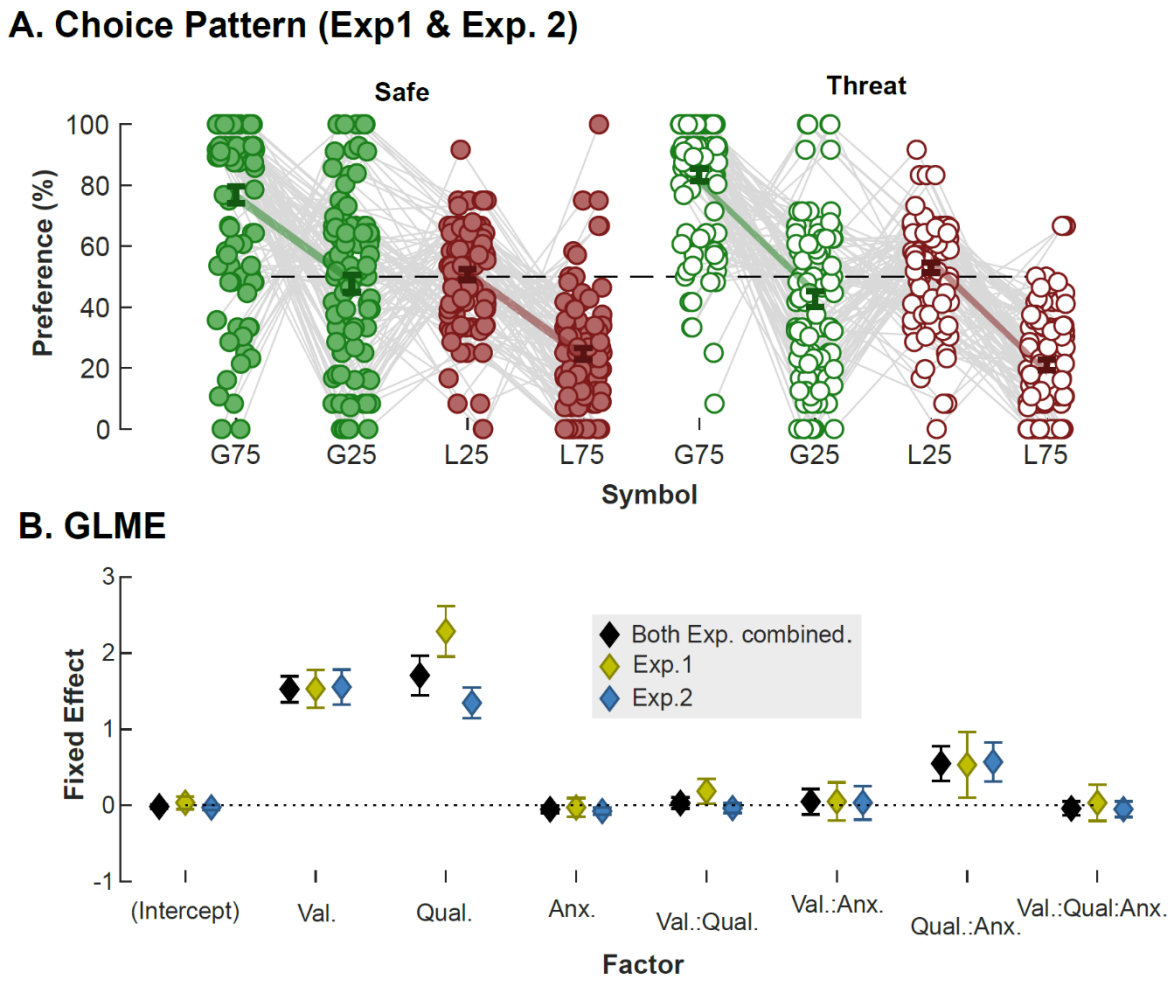


Figure 2.5. Choice pattern in transfer task and corresponding results from generalized linear mixed effect model. (A) Averaged choice rate for each cue. Each gray line indicates individual's choice pattern. The filled dots represent cues learned under the safe condition, the unfilled dots indicate cues learned during the threat condition. Error bars indicate the standard error of the mean (SEM), and shaded areas represent the mean ± SEM of the GLME predictions. **(B)** Generalized linear mixed-effect model (i.e., $GLME1_{Transfer}$) with cue selection as dependent variable. The y-axis represents the estimated standardized coefficient (t-value), and the x-axis represents each factor in the GLME model. Dot colors indicate results from different dataset.
G75: 75% of gain; G25: 25% of gain; L25: 25% of loss; G75: 75% of loss.
Val.: outcome valence; **Qual.**: Quality of cue (i.e. Higher expected value in its pair during learning); **Anx.**: Anxiety manipulation

Following previous studies (Palminteri et al., 2015; Palminteri, Lefebvre, et al., 2017), we conducted an additional analysis that focused on cues that were associated with intermediate values (i.e., G25 and L25) and tested if subject displayed rational preferences to choose cues based on expected value (**Table A.2.7**). The two-way ANOVA with valence and anxiety as factors showed

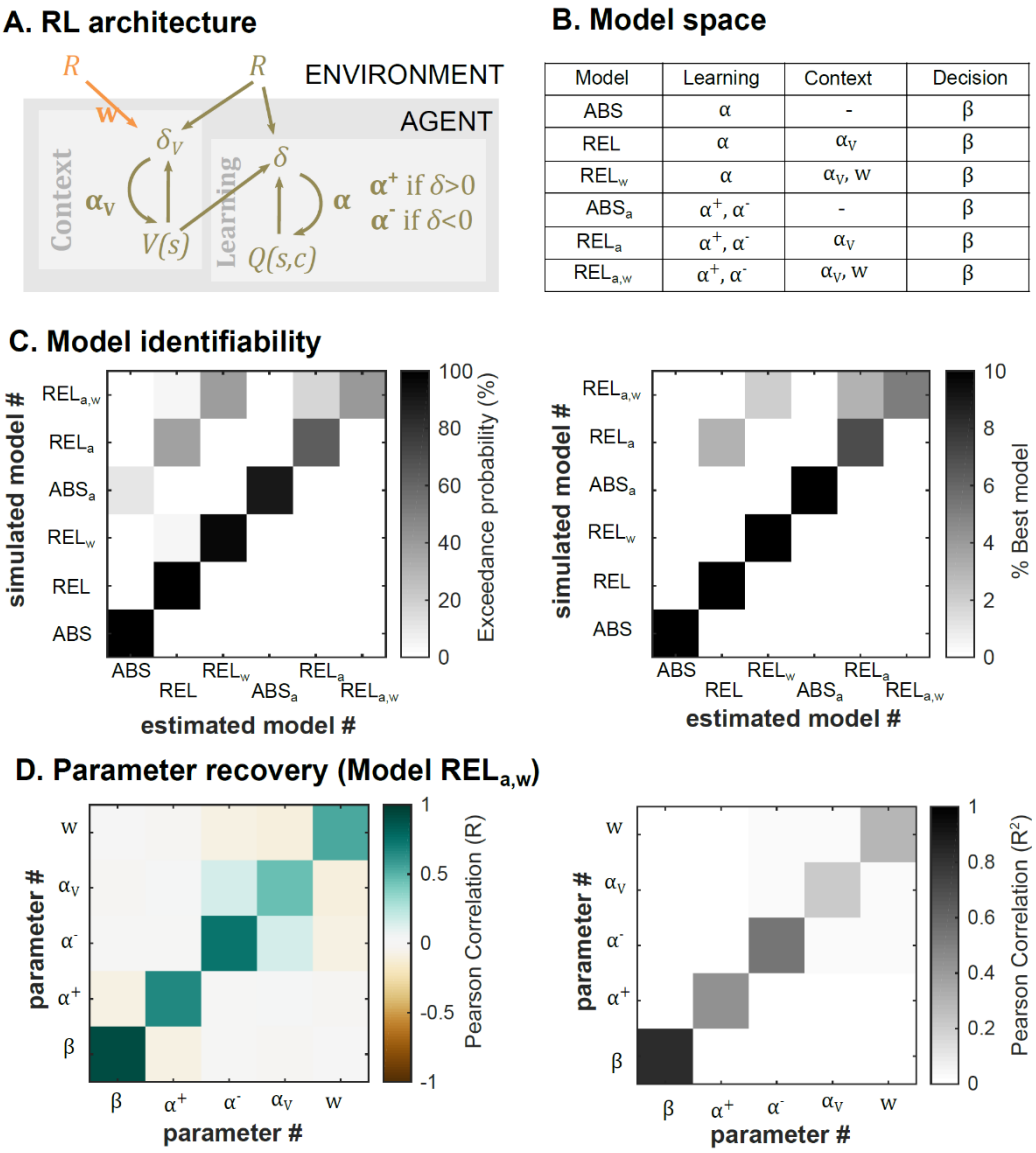
a significant main effect of valence ($F_{1,99}$ = 5.96, p = .0163, η_p^2 = 0.05), but no significant main effect of anxiety ($F_{1,99}$ = 0.55, p = .4586, η_p^2 = 0.00) nor its interaction ($F_{1,99}$ = 2.63, p = .1077, η_p^2 = 0.02).

2.3.4 Model-Based Analysis: learning is asymmetric and context-dependent

Our analysis of the general effects of anxiety in our factorial design (see **Section 2.2 Material and Methods** for details) points toward non-specific, elusive effects of anxiety in reinforcement learning. Two concurrent hypotheses might explain this observation: on the one hand, it is possible that anxiety affects specific latent mechanisms of reinforcement-learning, that may be subtle and therefore difficult to identify via model-free factorial design analyses; alternatively, it may be that anxiety indeed does not affect reinforcement-learning processes in a strong and idiosyncratic way.

To tease apart these competing explanations, we next turned to computational modelling. By explicitly modelling the computations giving rise to participants' behavior, computational modelling can efficiently combine data from both learning and transfer task (Palminteri et al., 2015) and can identify latent operations that would be specifically impacted by anxiety (Bishop & Gagne, 2018; Mkrtchian et al., 2017; Treadway et al., 2017). As a first step, we aimed at identifying a core architecture that would capture the learning behavior regardless of anxiety (i.e. in both safe and threat conditions). Following decades of research on the modelling of similar tasks (Pessiglione et al., 2006; Rescorla & Wagner, 1972; Sutton & Barto, 1998), we assumed that participants learned the value of available options using an algorithm akin to Q-learning.

Yet, several features of the observed behavioral pattern suggest that simple Q-learning would not be sufficient to comprehensively capture our participants' learning dynamics. First, the fact that participants generally express higher preference for the L25 than for the G25 cue in the transfer task is a signature of context-dependent learning (Klein et al., 2017; Lebreton et al., 2019; Palminteri et al., 2015; Palminteri, Lefebvre, et al., 2017). Briefly, in addition to standard Q-learning computations, context-dependent learning explicitly computes a context-value, which approximates the average expected value from a specific context. Obtained outcomes are then reframed relatively to this context value, allowing e.g. minor losses encountered in a loss context to be experienced as relative gains and vice-versa. This explains why small losses (L_{25}) are preferred to small gains (G_{25}) post learning in the transfer task. Second, the apparently higher variability of performance observed in the gain compared to the loss domain could be a signature of asymmetric learning. Briefly, if positive prediction errors are weighted more heavily than negative ones, individuals can quickly diverge in response rates (Lefebvre et al., 2017). Considering these two potential additional features of reinforcement learning models, we built a model space comprising six computational models presenting different combinations of those features (see **Section 2.2 Material and Methods** and **Figure 2.6 A-B**). Using simulations, we verified that those models were identifiable, and that their parameters could be satisfactorily estimated (**Figure 2.6 C-D**).



Model	Learning	Context	Decision
ABS	α	-	β
REL	α	α_V	β
REL_w	α	α_V, w	β
ABS_a	α^+, α^-	-	β
REL_a	α^+, α^-	α_V	β
$REL_{a,w}$	α^+, α^-	α_V, w	β

Figure 2.6. Modelling approach (Step 1). (A) depiction of the model architecture basis. **(B)** Model Space. **(C)** Model identifiability analysis. Data from 100 synthetic participants were simulated (50 with experiment 1 design, 50 with experiment 2 design) with each of our six models. Bayesian model selection was used to identify the most probable model generating the data, using the Laplace approximation to model evidence. This procedure was repeated 10 times. Left: average exceedance probability confusion matrix. Right: Best model selection confusion matrix. **(D)** Parameter recovery analysis. The confusion matrices represent summary statistics of the correlations between parameters, estimated over 100-subject simulations, and averaged over the 10 simulations. Diagonal: correlations between simulated and estimated parameters. Off diagonal: cross correlation between estimated parameters. Left: Pearson correlation (R). Right: explained variance (R^2).

We then fitted those models to our data, and ran a full Bayesian model comparison (BMC) procedure, aiming at identifying the best and most parsimonious computational architecture. In line with the behavioral signatures identified in the beginning of this section, the BMC identified the RL model including both learning-rate asymmetry and context-dependency ($REL_{a,w}$) as the best explanation of our data (exceedance probability: 93%; **Figure 2.7A**). Note that $REL_{a,w}$ also won the BMC procedure based on data restricted to experiment 1 and 2 (exceedance probability: 76% and 88%; **Figure 2.7A**). Average estimated parameter values were very similar to previous studies (Palminteri et al., 2015), and were also very similar between experiment 1 and 2 (**Figure 2.7B**). The modelling results notably replicate the learning asymmetry reported in previous studies (Lefebvre et al., 2017; Palminteri, Lefebvre, et al., 2017), with positive learning rates being significantly larger than negative learning rates ($\alpha^+ = 0.37 \pm 0.02$; $\alpha^- = 0.07 \pm 0.01$; $t_{99} = 10.26$, $p = 2.92 \times 10^{-17}$, $d = 1.06$). Overall, this model provided a very good fit of both learning and transfer task data (**Figure 2.7 C-D**).

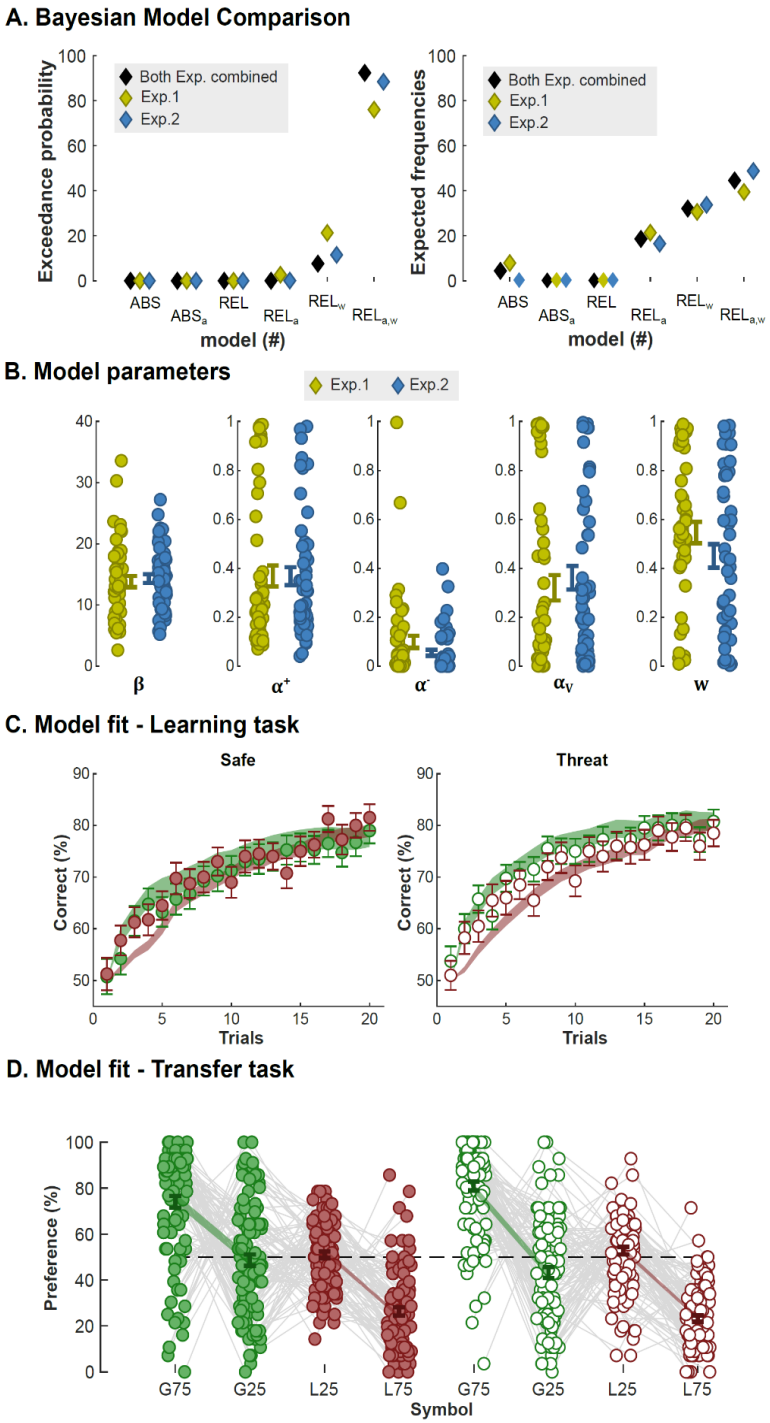


Figure 2.7. Modelling results for the general learning architecture (Step1). (A) Model comparison results. These panels depict the results of a Bayesian model comparison analysis on our participants data, for experiment 1 (yellow diamonds), experiment 2 (blue diamonds) and both experiments combined (black diamonds). Left: exceedance probability of each model. Right: expected frequencies of each model. **(B)** Model parameters of the winning model ($REL_{a,w}$) for experiments 1 (yellow) and 2 (blue). Filled dots represent individual parameters while error bars represent population mean±SEM. **(C)** Learning curves, representing the fraction of correct choices in the safe (left; filled dots) and the threat (middle; unfilled dots) conditions. Error bars indicate the standard error of the mean (SEM), and shaded areas represent the mean ± SEM of the $REL_{a,w}$ predictions. **(D)** Transfer choice rate for each cue. Each gray line indicates individual's choice pattern. The filled dots represent cues learned under the safe condition, the unfilled dots indicate cues learned during the threat condition. Error bars indicate the standard error of the mean (SEM), and shaded areas represent the mean ± SEM of the $REL_{a,w}$ predictions.

2.3.5 Model-Based Analysis: the effects of anxiety on learning is inconclusive

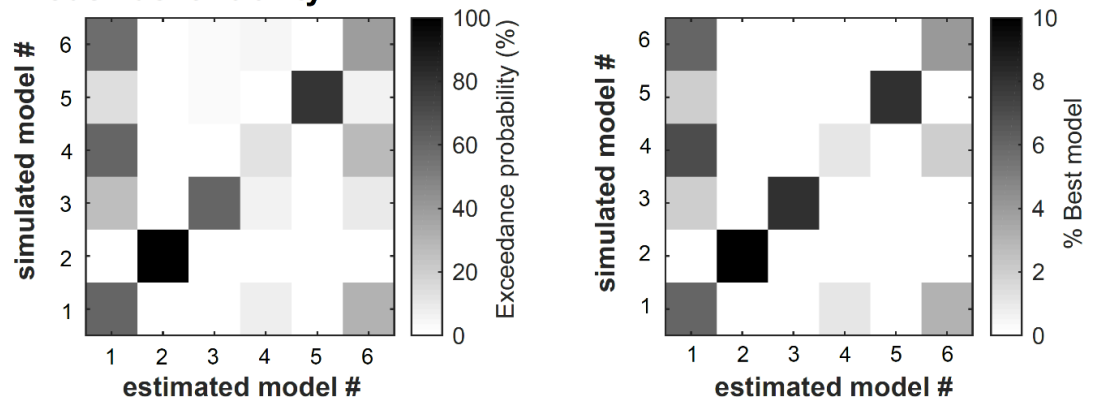
To investigate the effects of anxiety on learning, we next built a second model space, where all models were derived from the $REL_{a,w}$, but additionally allowed parameters to vary between the Safe and Threat conditions. Given that $REL_{a,w}$ possesses five parameters (the choice temperature, three learning rates, and the unchosen outcome weighting parameter), the second model space featured six models (see **Figure 2.6B**). Similar to the first modelling step, we ran model identifiability and parameter recovery analyses (**Figure 2.8B**). Results show that some models cannot be perfectly identified: models 4 and 6, which respectively feature differential learning rates for negative PE (α^-) and differential weighting parameters (w) between anxiety and safe conditions both tended to be identified as the simplest model.

Despite those limitations, we still compared those models in their ability to account for the observed data (see **Section 2.2.5 Material and Methods – Computational Modelling**). The Bayesian Model Comparison with LAME failed to identify a clear best model (**Figure 2.8C**), indicating that allowing important model parameters to vary as a function of anxiety does not improve model fit.

A. Model space

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
β	S = T	S ≠ T	S = T	S = T	S = T	S = T
α^+	S = T	S = T	S ≠ T	S = T	S = T	S = T
α^-	S = T	S = T	S = T	S ≠ T	S = T	S = T
α_V	S = T	S = T	S = T	S = T	S ≠ T	S = T
w	S = T	S = T	S = T	S = T	S = T	S ≠ T

B. Model identifiability



C. Bayesian Model Comparison

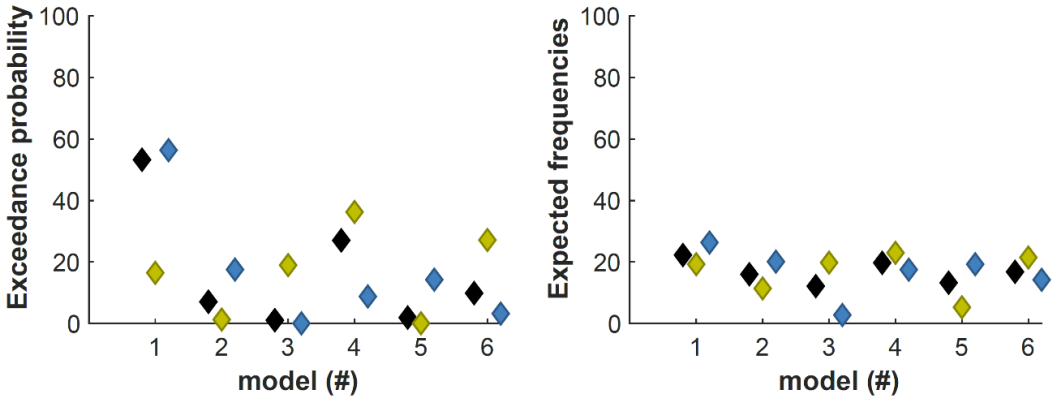


Figure 2.8. Modelling approach (Step 2). (A) Model Space. S and T represent safe and threat condition, respectively. S = T indicates safe and threat condition share the same parameter. **(B)** Model identifiability analysis. Data from 100 synthetic participants were simulated (50 with experiment 1 design, 50 with experiment 2 design) with each of our six models. Bayesian model selection was used to identify the most probable model generating the data, using the Laplace approximation to model evidence. This procedure was repeated 10 times. Left: average exceedance probability confusion matrix. Right: Best model selection confusion matrix. **(C)** Model comparison results. These panels depict the results of a Bayesian model comparison analysis on our participants data, for experiment 1 (yellow diamonds), experiment 2 (blue diamonds) and both experiments combined (black diamonds). Left: exceedance probability of each model. Right: expected frequencies of each model.

2.4 Discussion

2.4.1 Summary of design, results and contributions

In two experiments, we investigated the impact of incidental anxiety (i.e., unpredictable threat) on both learning performance during a probabilistic instrumental task (Palminteri et al., 2015; Pessiglione et al., 2006) and on post-learning transfer preferences (Frank et al., 2004; Palminteri et al., 2015), using a well-established Threat of Shock (ToS) paradigm (Cohn et al., 2015; Engelmann et al., 2015, 2019; Grillon, 2008; Schmitz & Grillon, 2012). Two variants of the anxiety induction method were used during learning: threat and safety trials were either alternated in blocks of three trials (experiment 1), which induced relatively higher levels of anxiety, or in sessions of 80 trials (experiment 2), which induced relatively lower levels of anxiety. Behavioral results from two experiments consistently showed that anxiety and outcome valence had little to no effect on learning performance *per se*. At first glance, these results may be somewhat surprising given that they seem to contradict several previous studies suggesting that anxiety alters learning performance (DeVido et al., 2009; Schwabe & Wolf, 2009; Stevens et al., 2014; Treadway et al., 2017). However, a small-scale literature review agrees with our main result as it also failed to identify consensual and robust main effects of anxiety on learning across 13 papers (**Figure 2.1** and **Table A.1**).

Importantly, our results nonetheless suggest that transfer preferences were significantly altered by the safe/anxiety manipulation. Specifically, post-learning preferences indicate that participants were better able to identify the quality of cues when these cues were learned in a threatening compared to a safe context. Note again that this effect was observed in the absence of any effects of anxiety on *average learning performance* and therefore is indicative of anxiety exhibiting somewhat delayed effects on post-learning preferences and recall (see also **Section 2.4.3 Discussion: transfer task**). Similar improvements in the ability to identify cues during a post-learning transfer task have been observed in one other prior study (Cavanagh et al., 2011).

Considering the possibility that anxiety effects differ when individuals seek rewards vs. avoid punishments, our experiment did not reveal any valence-dependent effects during both learning and post-learning performance. These results agree with our targeted literature review, which also failed to identify a robust consensus on this question. Specifically, other studies have reported inconsistent valence dependent effects of threat states and trait anxiety on post-learning performance (Abraham and Hermann, 2015; Berghorst et al., 2013; Cavanagh et al., 2019, 2011; Lighthall et al., 2013; Petzold et al., 2010; Voegler et al., 2019): for learning performance in the domain of gains 6/8 studies reported a null effect, 1/8 reported improvements and 1/8 reported reduced performance. In the domain of losses, 4/8 studies reported improved performance, while 1/8 reported reduced performance in the domain of losses and 3/8 reported a null effect (**Figure 2.1**). Jointly, the current and previous results indicate that anxiety likely does not have differential effects on learning to seek rewards and to learning to avoiding punishments.

A primary goal of the current experiments was to assess multiple experimental factors that could skew results from prior experiments in one experimental setup. To this end, our experiments were carefully designed to assess the differential effects of anxiety on learning and post-learning preferences, as well as on punishment and reward learning, while simultaneously reducing the effects of potential confounding factors using multiple methods. Firstly, by adapting a ToS

paradigm to induce anxiety, each subject learned action-outcome associations separately under a threatening context and under safety. Moreover, the ToS procedure allowed us to customize the intensity of the negative event (i.e., electric shock) to each participant's pain threshold, and to successfully create significant threat levels for all participants across two experiments (as assessed by SCR responses and self-reports). The ability to turn threat on and off at specific time points throughout the experiments allowed us to directly assess the effects of anxious states on learning in a within-subject design. This is important to assess the causal effects of anxiety on learning (Engelmann et al., 2015, 2019), and addresses a major limitation of traditional emotion and stress induction techniques, the common delay between the induction time point and behavioral task. Secondly, we induced anxiety for two different periods including relatively short blocks of three trials (experiment 1) and relatively long periods lasting for full session of 96 trials (experiment 2). This was done for two reasons; (1) it allowed us to assess potential biases induced by the repeated switching of emotional states and associated stimuli in experiment 1, and (2) it allowed us to assess the effect of different threat level intensities, with experiment 1 creating a relatively more intense emotional state compared to experiment 2. Thirdly, we assessed the effects of anxiety on learning over gains and losses by crossing the ToS manipulation with an outcome valence manipulation. This allowed us to directly assess the effects of anxiety on learning to seek gains and to avoid losses separately (Palminteri & Pessiglione, 2017). Finally, we differentiated the effects of anxiety on learning and post-learning performance by including both a learning stage and a post-learning task in the same study. Despite these methodological advances, we find only limited effects of anxiety on learning per se, but significant enhancements of the ability to identify better quality cues in a post learning transfer task.

2.4.2 Discussion: learning

We consider three explanations for the limited effects of incidental anxiety on learning, observed in the current two experiments. These include the hypotheses that (1) only trait, but not state anxiety may have an impact on learning; (2) anxiety reduces the available cognitive resources and its effects can only be revealed in more difficult settings; (3) anxiety causes an inflexibility in learning, which can only be revealed using more complex task designs that require planning and adaptation.

Firstly, a prominent hypothesis suggests that state and trait anxiety might impact different underlying learning mechanisms (Robinson, Vytal, et al., 2013; Robinson & Chase, 2017). Specifically, Robinson et al argue that state anxiety in response to unpredictable threat is "adaptive" as it prepares flexible physiological responses and behaviors to cope with negative events that may ensure survival (Robinson et al., 2013). Elevated and prolonged levels of trait anxiety on the other hand can become "maladaptive" and interfere with normal day-to-day functioning (Robinson et al., 2013). A number of previous studies showed behavioral effects of trait and pathological anxiety on learning that generally support this hypothesis (**Figure 2.1**, yellow bars; Abraham & Hermann, 2015; Browning et al., 2015; DeVido et al., 2009; Robinson et al., 2015; Stevens et al., 2014; Voegler et al., 2019). However, this idea has recently been challenged by research revealing more limited effects of both trait and state anxiety on learning performance (Berghorst et al., 2013; Cavanagh et al., 2011; Lighthall et al., 2013; Petzold et al., 2010; Schwabe & Wolf, 2009; Voegler et al., 2019),

even in more complicated settings (Two-stage reinforcement learning: Gillan et al., 2020; social learning: Safra et al., 2018). To resolve this disagreement on the relative importance of state compared to trait anxiety, we assessed the effects of trait anxiety in our current data set by conducting additional analyses focusing on individual differences in BAI scores on learning and post-learning performance (see **Appendix A.4**). Note that our approach allows us to identify both main effects of state and trait anxiety, but also their interaction. Our results did not identify any main effects of BAI in the learning and post-learning task, indicating that high and low anxious participants learned to seek gains and avoid losses equally well. Moreover, we did not observe an interaction effect between state and trait anxiety during learning and post-learning. Our results, together with the inconsistent findings of prior research identified by our literature review, suggest that both trait and state anxiety have little main and interactive effects on learning and post-learning performance.

Secondly, according to attentional control theory (Eysenck et al., 2007), performance might be intact if cognitive demands of the task do not exceed cognitive resources. Therefore, if the probability of a positive or negative outcome differs substantially between the stimuli during the learning phase (i.e., 75% vs 25%), the task might be too easy to reveal the impact of anxiety given the experimental setup in current study. Specifically, participants might have sufficient cognitive resources to deal with such a relatively simple task, even when learning under conditions of anxiety. Although this explanation is partially rejected by previous studies that found no significant interactions between difficulty and neither threat manipulations neither trait anxiety, the tasks used in previous experiments might also not have sufficiently challenged participants (Abraham & Hermann, 2015; Berghorst et al., 2013; Cavanagh et al., 2011, 2019; Lighthall et al., 2013; Petzold et al., 2010; Voegler et al., 2019). To address this potential explanation, we inspected our data by focusing on a subset of participants that showed evidence for finding this task relatively difficult. We identified these participants via cluster analysis (k-means, **Appendix A.5**) and split our subject pool into two groups, one showing relatively lower average performance (57% accuracy) throughout the learning task and another with relatively higher average performance (76% accuracy). If the predictions from attentional control theory apply to our results, we should observe larger effects of anxiety on learning and post-learning performance in the subject group with lower average performance that likely found the task more difficult. Our results did not support this potential interpretation and were consistent with previous findings (Abraham & Hermann, 2015; Berghorst et al., 2013; Cavanagh et al., 2011, 2019; Lighthall et al., 2013; Petzold et al., 2010; Voegler et al., 2019; but see Robinson et al., 2013 for the impact of anxiety on prediction error processing in the brain). Specifically, while we found a significant main effect of performance group on predicting correct choice in the learning task, this effect was not modulated by anxiety (see **Appendix A.5**). The result in the transfer task also showed non-significant two- and three-way interactions of anxiety with performance group and quality. These results further support the notion that anxiety has limited effects on learning performance. Moreover, the effects we observe here are likely not modulated by task-dependent availability of cognitive resources (see also Engelmann et al., 2015).

In a similar vein, the anxiety condition might specifically increase cognitive task load. We inspected this possibility by analyzing reaction times across the two conditions, which is generally considered one of the hallmark measures of cognitive load (Pashler, 1994). We did not find slower reaction

times under conditions of anxiety (**Table A.2.9**), indicating that participants did not face greater cognitive load in the threat condition.

Finally, the effects of anxiety have previously been associated with inflexibility in learning, which might be caused by either difficulty in switching between habit and goal-directed decisions (Browning et al., 2015; Otto et al., 2013; Raio et al., 2017; Schwabe & Wolf, 2009) or intolerance for uncertainty (Behar et al., 2009; Buhr & Dugas, 2006; Dugas et al., 1995). Therefore, our one-stage reinforcement learning task with stable contingency of action and outcome might not be able to detect these anxiety-related changes in behavior. Taken together, the impact of anxiety on learning might be varied and depend on the type of anxiety and the task's difficulty and their interactions. Yet, again, the robustness of the effects of anxiety on learning flexibility have been challenged by recent high-powered studies (Gillan et al., 2020)[12] and our small-scale review.

2.4.3 Discussion: transfer task

The significant general improvement in the ability to identify better options during the transfer task when these were learned under anxiety is consistent with a growing literature showing that a threatening environment can significantly impact memory processes under specific conditions (Bolton & Robinson, 2017; Vytal et al., 2013). In light of this, the current results might indicate that anxiety enhances memory retrieval for the value of cues encoded under anxiety (Mather & Lighthall, 2012; Porcelli & Delgado, 2017).

Note that the presence of anxiety effects observed in the transfer task in the absence of any effects of anxiety on *average learning performance* could also indicate that the transfer task is more sensitive to capture anxiety effects, whereas average learning effects dilute them. We tested this possibility, by analyzing separately the early and late phase of learning (see **Appendix A.7**). Our results suggest that the learning performance in the late learning stage, rather than early or overall performance, might be more susceptible to the effect of incidental anxiety – although these effects are still very marginal and would need to be replicated.

The absence of detectable valence-specific effects of anxiety on performance in the transfer task in the present study might at first glance contradict previous results on the effects of anxiety on learning (Berghorst et al., 2013; Cavanagh et al., 2019; Lighthall et al., 2013; Petzold et al., 2010; Voegler et al., 2019). However, our targeted literature review, suggests that prior results are rather inconsistent with no detectable trends in the domain of gains (null effect: 75%, improvement: 12.5%, decline: 12.5%) and suggest a slight improvement in average performance in the domain of losses (null effect: 37.5%, improvement: 50%, decline: 12.5%).

[12] Gillan et al., 2020 conducted both correlational and causal experiments.

2.4.4 Discussion: Modelling

Besides model-free analyses, we also used computational modeling as a more formal tool to characterize the specific influence(s) of anxiety on the learning process. Our modelling approach, validated by simulation-based parameter recovery and model indentifiability procedures, identified a winning model that updates expected values using a context-dependent and asymmetric learning rule (Palminteri et al., 2015; Wilson & Collins, 2019). Context-dependency implies that option values are updated with respect to a reference point, which approximate the average expected value of a pair, and which is learned on a trial-by-trial basis with a specific contextual-learning rate (Palminteri et al., 2015). Factually, this allows to reframe small losses as rewards in a loss context and small rewards as losses in a gain context. Learning asymmetry was featured by different learning rates to update values after positive vs negative prediction errors. Similarly to previous reports, we found that learning asymmetry captures a confirmation bias, with positive learning rate parameters taking values twice as big as negative learning rate parameters (Lefebvre et al., 2017; Palminteri, Lefebvre, et al., 2017).

These results have important implications: they suggest that human learning incorporates more features (context-dependency, learning asymmetry) than typically thought, even in simple, traditional instrumental learning tasks that have been used for years (Frank et al., 2004; Pessiglione et al., 2006). Because interpreting parameter fits from models that provide incomplete descriptions of behavior is problematic (Nassar & Gold, 2013), this suggests that some simple modelling approaches that omit those features could have converged on erroneous conclusions about the effects of experimental manipulations or neuro-psychiatric pathologies on learning parameters.

After having identified the model that best and most comprehensively accounts for the general learning behavior of our participants, we aimed to evaluate the impact of anxiety on its parameters. Yet, we found that models including extra parameters to capture the effect of anxiety cannot be robustly identified. This misidentification issue suggests that those parameters have such a subtle (i.e., small) effect on the general behavior observed in the learning and transfer tasks, that the current task design (with its conditions, number of sessions, and number of trials) is not powerful enough to detect them. Given that our experimental design favorably compares to previous ones in terms of power (number of subjects, trials, etc.), this indicates that most designs (including ours) might not be powerful enough to allow the detection of the potential effects of anxiety, once all the complex features of learning that can be detected in human learning behavior (context dependency, learning asymmetry) are taken into account.

2.4.5 Conclusion

The current study investigated the effects of anxiety on learning, while addressing several concerns about experimental designs and analytical choices that might have led to discrepancies in the identification of such effects in previous studies. Despite our relatively powerful approach that simultaneously assessed learning and post-learning performance, as well as reward and punishment learning in the context of a within-subject anxiety induction, and contrary to some previous studies, our experiments failed to reveal clear and specific effects of anxiety on learning per se. While surprising at first glance, our null results agree with findings from a small-scale review that shows little to no effects of anxiety on learning and post-learning performance on average and they add to recent results, which have started to challenge the role of anxiety in experience-based decision-making (Bishop & Gagne, 2018; Gillan et al., 2020).

CHAPTER 3

Robust Valence-Induced Biases on Motor Response and Confidence in Human Reinforcement Learning[13]

[13] This chapter is based on Ting, C-C., Palminteri, S., Engelmann, J.B., Lebreton, M. (2020b) Robust valence-induced biases on motor response and confidence in human reinforcement learning. Cognitive, Affective, & Behavioral Neuroscience.

3.1 Introduction

In the reinforcement learning context, reward-seeking and punishment-avoidance present an intrinsic and fundamental informational asymmetry. In the former situation, accurate choice (i.e., reward maximization) increases the frequency of the reinforcer (the reward). In the latter situation, accurate choice (i.e., successful avoidance), optimal behavior decreases the frequency of the response. Accordingly, any simple incremental "law-of-effect"-like model, would predict higher performance in the reward seeking compared the punishment avoidance situation. Yet, our study (**Chapter 2;** Ting et al., 2020a) and previous studies (Fontanesi et al., 2019; Guitart-Masip et al., 2012; Palminteri et al., 2015) demonstrated that humans learn to seek reward and to avoid punishment equally-well. This is not only robustly demonstrated in experimental data, but also nicely explained by context-dependent reinforcement-learning models (Fontanesi et al., 2019; Palminteri et al., 2015), which can be seen as formal computational instantiation of Mowrer's two-factor theory (Mowrer, 1952). On top of this remarkable symmetry in choice accuracy between gain and loss contexts, two sets of recent studies independently reported that outcome valence asymmetrically affects confidence and response times (RTs). First, learning from punishment increases individuals' RTs, slowing down the motor execution of the choice (Fontanesi et al., 2019; Jahfari et al., 2019; Ting et al., 2020a). This robust phenomenon is consistent with a motor Pavlovian bias which posits that desirable contexts favor motor execution and approach behavior, while undesirable contexts hinder them (Boureau & Dayan, 2011; Guitart-Masip et al., 2012).

Second, learning from punishment decreases individuals' confidence in their choices (Lebreton et al., 2019). Confidence judgements can be defined and operationalized as the subjective estimations of the probability of being correct (Fleming & Daw, 2017; Pouget et al., 2016). As such, a confidence judgment is a metacognitive operation, which quantifies the degree to which an individual is aware of his or her success or failure (Fleming & Dolan, 2012; Yeung & Summerfield, 2012). Confidence judgments are thought to be critical in the context of meta-control – the flexible adjustment of behavior –, as they are key to monitor and reevaluate previous decisions (Folke et al., 2016) to track changes in the environment (Heilbron & Meyniel, 2019; Vinckier et al., 2016), or to arbitrate between different strategies (Daw et al., 2005; Donoso et al., 2014). The demonstrations that confidence judgments can be biased by the outcome valence in different tasks (Lebreton et al., 2018, 2019) suggest that, similarly to instrumental processes, metacognitive processes could also be under the influence of Pavlovian processes.

Here, we aimed to investigate the link between the valence-induced motor and confidence biases. We focused on two research questions: first, are valence-induced motor and confidence biases robust and replicable? Second, can the confidence bias be observed in the absence of the motor bias? Regarding the second question, previous research has yielded conflicting results that generated two opposing predictions. On the one hand, numerous studies documented behavioral and neural dissociations between perceptual, cognitive or motor operations, and confidence or metacognitive judgments (Fleming et al., 2012; Miele et al., 2011; Qiu et al., 2018). Likewise, brain lesions and stimulation protocols have been shown to disrupt confidence ratings and metacognitive abilities without impairing cognitive or motor functions (Fleming et al., 2014, 2015; Rounis et al., 2010) - although see also (Bor et al., 2017). These dissociations between decision and metacognitive variables suggest that the valence-induced confidence bias could be observed in the

absence of a response time bias.

On the other hand, several studies suggest that decision and metacognitive variables are tightly linked– both in perceptual (Geller & Whitman, 1973; Vickers et al., 1985) and value-based tasks (De Martino et al., 2013; Folke et al., 2016; Lebreton et al., 2015). This coupling is notably embedded in many sequential-sampling models which rely on a single mechanism to produce decisions, response times and confidence judgments (van den Berg et al., 2016; De Martino et al., 2013; Moran et al., 2015; Pleskac and Busemeyer, 2010; Ratcliff and Starns, 2009, 2013; Yu et al., 2015). Beyond this mechanistic hypothesis, it was also recently suggested that people use their own RT as a proxy for stimulus strength and certainty judgments, creating a direct, causal link from RT to confidence (Desender, Opstal, et al., 2017; Kiani et al., 2014). These results could imply that our previously reported effects of valence on confidence (Lebreton et al., 2019) are no more than a spurious consequence of the effect of valence on RTs (Fontanesi et al., 2019; Jahfari et al., 2019). In other words, participants could have simply observed that they were slower in the loss context, and used this information to generate lower confidence judgments in these contexts.

In order to address our research questions, we developed several versions of a probabilistic, instrumental-learning task, where participants have to learn to seek rewards or to avoid losses (Fontanesi et al., 2019; Lebreton et al., 2019; Palminteri et al., 2015). We attempted to cancel the effects of losses on RTs while recording confidence judgments to assess the presence of the valence-induced confidence bias. To this end, we modified the standard mapping between the available options and the way participants could select them, thereby disrupting the link between decision and motor execution of the choice. In another experiment, we also used a different strategy, and imposed time pressure on the choice to constrain decision time.

In total, we used two published datasets (Lebreton et al., 2019) and original data collected from four new experiments, where we manipulated in several ways the option-action mapping (experiment 3-5) and applied time pressure (experiment 6). We then tested (1) the robustness of the valence-induced motor and confidence biases; (2) whether the confidence bias could be observed in the absence of the motor bias. Overall, our results show that response times are slower in loss than gain contexts in almost all experiments. In other words, the motor bias is highly robust, as it survived most of our disruption attempts, despite being severely attenuated. In all datasets, confidence was lower in loss than in gain contexts, indicating that the confidence bias is highly replicable, and is robust to variations in the motor bias effect sizes. The confidence bias is also observed in the condition where the motor bias was absent, suggesting that valence-induced motor and confidence biases are – partly – dissociable.

3.2 Material and Methods

3.2.1 Participants

All studies were approved by the local Ethics Committee: Economics and Business Ethics Committee (EBEC), at the University of Amsterdam. All participants gave informed consent prior to partaking in the study. The participants were recruited from the laboratory's participant database (www.creedexperiment.nl). A total of 108 participants took part in this set of 6 separate experiments (**Table 3.1**). They were compensated with a combination of a show- up fee (5€), and additional gains and/or losses depending on their performance during the learning task: experiment 1 had an exchange rate of 1 (in-game euros = payout); experiments 2-6 had an exchange rate of 0.3 (in game euros = 0.3 payout euros, participants were clearly informed of this exchange rate). In addition, in experiments 2-6, three trials (one per session) were randomly selected for a potential 5 euros bonus each, attributed based on the confidence incentivization scheme (see below).

3.2.2 Power analysis and sample size determination. Power analysis were performed with GPower.3.1.9.2, and followed the reasoning in (Lebreton et al., 2019). The sample size for all experiments was determined prior to the start of the experiments based on the effects of incentives on confidence judgments in (Lebreton et al., 2018). Cohen's d was estimated from a GLM d = .941 $t_{23} = 4.61$, $p = 1.23 \times 10^{-4}$. For a similar within-subject design, a sample of N = 17 subjects was required to reach a power of 95% with a two-tailed one-sample t-test.

3.2.3 Learning task - General

In this study, we iteratively designed six experiments, aiming at investigating the impact of context valence and information on choice accuracy, confidence and response times, in a reinforcement-learning task. All experiments were adapted from the same basic experimental paradigm (**Figure 3.1A** and see also **Appendix B.1** for details): participants repeatedly faced pairs of abstract symbols probabilistically associated with monetary outcomes (gains or losses), and they had to learn to choose the most advantageous symbol of each pair (also referred to as context), by trial and error. Two main factors were orthogonally manipulated (Palminteri et al., 2015): valence (i.e. some contexts only provide gains, and others losses) and information (some contexts provide information about the outcome associated with both chosen and unchosen options –complete information- while others only provided information about the chosen option –partial information). In addition, at each trial, participants reported their confidence in their choice on a graded scale as the subjective probability of having made a correct choice (**Figure 3.1A**). In all experiments but one (Exp. 2-6) those confidence judgments were elicited in an incentive-compatible way (Ducharme and Donnell, 1973; Lebreton et al., 2018, 2019; Schlag et al., 2015).

Results from experiment 1 and 2 were previously reported in (Lebreton et al., 2019): briefly, we found that participants exhibit the same level of choice accuracy in gain and loss contexts, but are less confident in loss contexts. In addition, they appeared to be slower to execute their choices in loss contexts. Here, in order to evaluate the interdependence between the effects of valence on RT and confidence, we successively designed three additional tasks (**Figure 3.1A**; see also

Appendix B.1 for details). In those tasks, we modified the response setting to blur the effects of valence on RT, with the goal to assess the effects of valence on confidence in the absence of an effect on RT. In a sixth task we imposed a strict time pressure on decisions (**Figure 3.1A; Figure B.1**). All participants also performed a Transfer task (Lebreton et al., 2019; Palminteri et al., 2015). Data from this additional task is not relevant for our main question of interest and is therefore not analyzed in the present manuscript.

3.2.4 Learning task - Details

All tasks were implemented using MatlabR2015a® (MathWorks) and the COGENT toolbox (http://www.vislab.ucl.ac.uk/cogent.php). In all experiments, the main learning task was adapted from a probabilistic instrumental learning task used in a previous study (Palminteri et al., 2015). Invited participants were first provided with written instructions, which were reformulated orally if necessary. They were explained that the aim of the task was to maximize their payoff and that gain seeking and loss avoidance were equally important. In each of the three learning sessions, participants repeatedly faced four pairs of cues - taken from Agathodaimon alphabet. The four cue pairs corresponded to four conditions, and were presented 24 times in a pseudo-randomized and unpredictable manner to the subject (intermixed design). Of the four conditions, two corresponded to reward conditions, and two to loss conditions (**Figure 3.1 B-C**). Within each pair, and depending on the condition, the two cues of a pair were associated with two possible outcomes (1€/0€ for the gain and -1€/0€ for the loss conditions in Exp. 1; 1€/0.1€ for the gain and -1€/-0.1€ for the loss conditions in Exp. 2-6) with reciprocal (but independent) probabilities (75%/25% and 25%/75%) - see Lebreton et al., 2019 for a detailed rationale.

Experiments 1, 2 and 6 were very similar (**Figure 3.1A, top**): at each trial, participants first viewed a central fixation cross (500-1500ms). Then, the two cues of a pair were presented on each side of this central cross. Note that the side in which a given cue of a pair was presented (left or right of a central fixation cross) was pseudo-randomized, such as a given cue was presented an equal number of times on the left and the right of the screen. Participants were required to select between the two cues by pressing the left or right arrow on the computer keyboard, within a 3000ms (Exp. 1-2) or 1000ms (Exp. 6) time window. After the choice window, a red pointer appeared below the selected cue for 500ms. Subsequently, participants were asked to indicate how confident they were in their choice. In Experiment 1, confidence ratings were simply given on a rating scale without any additional incentivization. To perform this rating, they could move a cursor –which appeared at a random position- to the left or to the right using the left and right arrows, and validate their final answer with the spacebar. This rating step was self-paced. Finally, an outcome screen displayed the outcome associated with the selected cue, accompanied with the outcome of the unselected cue if the pair was associated with a complete-feedback condition.

In experiment 3, we dissociated the option display and motor response: symbols were first presented on a vertical axis (2s), during this period, participants could choose their preferred symbol, but were uncertain about which button to press to select their preferred symbol. This uncertainty was resolved in the next task phase, in which two horizontal cues indicated which of the left vs right response button could be used to select the top vs bottom symbol (**Figure B.1C**). In addition, we imposed a time limit on the response selection (<1s), to incentivize participants to make their decision during the symbol presentation, and allow only an execution of a choice that

was already made during the response mapping screen. In Experiment 4, we added a mask (empty screen 0.5-1s) between the symbol presentation and the response mapping (**Figure B.1D**). This further strengthened the encouragement to make a decision during the symbol presentation to reduce task load, because participants would then only have to retain the information about the selected location (top vs bottom) during the mask period. In Experiment 5, we introduced a jitter (variable time duration; 2-3s) at the symbol presentation screen (**Figure B.1E**) to further discourage temporal expectations and motor preparedness during the decision period. Finally, Experiment 6 was adapted from Experiment 2, but additionally imposed a strict time pressure on the choice, in an attempt to incentive participants to counteract the slowing down due to the presence of losses (**Figure B.1F**). In all experiments, response time is defined as the time between the onset of the screen conveying the response mapping (Symbol for Exp. 1-2 & 6; Choice for Exp. 3-5; see **Figure 3.1A; Figure B.1**), and the key press by the participant.

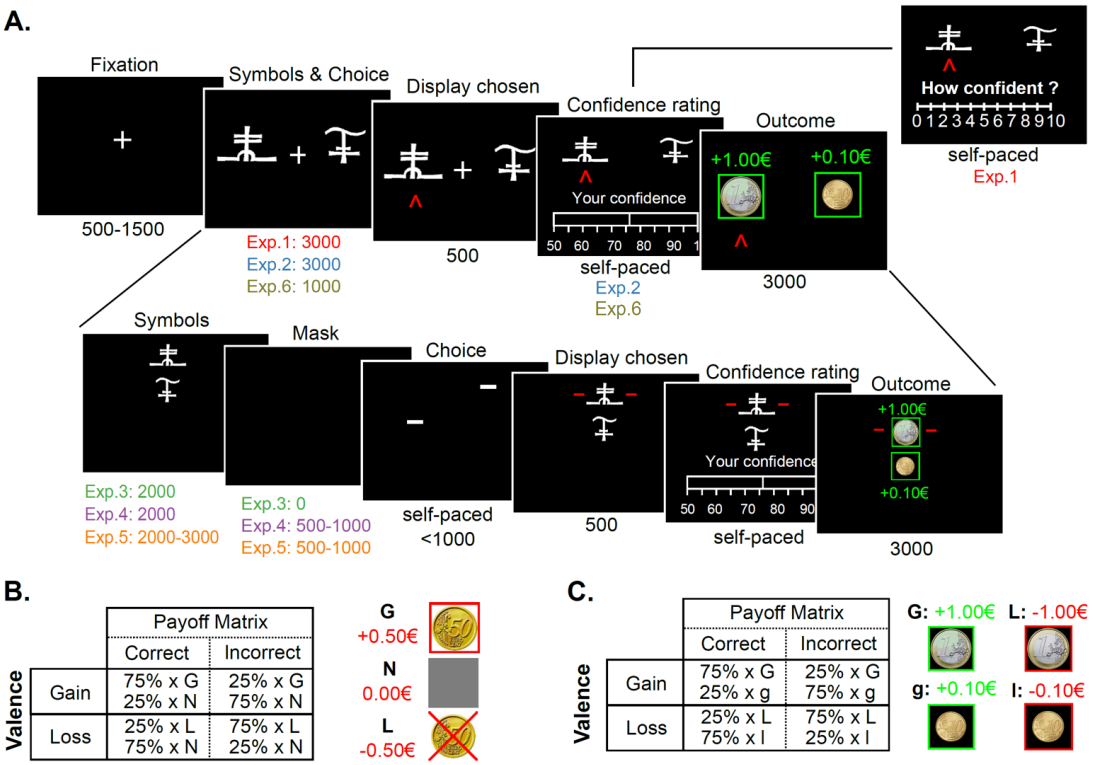


Figure 3.1 Experimental paradigms. (A) Behavioral tasks for Experiments 1-6. Successive screens displayed in one trial are shown from left to right with durations in milliseconds (ms). All tasks are based on the same principle, originally designed for experiments 1-2 (top line): after a fixation cross, participants are presented with a couple of abstract symbols displayed on a computer screen and have to choose between them. They are thereafter asked to report their confidence in their choice on a numerical scale. Note that experiment 1 featured a 0-10 scale, and experiments 2-6 featured a 50-100% scale. Outcome associated with the chosen symbol is revealed, sometimes paired with the outcome associated with the unchosen symbol -depending on the condition. For experiments 3-5 (bottom line), options are displayed on a vertical axis. Besides, the response mapping (how the left vs right arrow map to the upper vs lower symbol) is only presented after the symbol display, and the response has to be given within one second of the response mapping screen onset. A short empty screen is used as a mask, between the symbol display and the response mapping for experiments 4-5. Experiment 6 is similar to experiment 2 (top line), except that a shorter duration is allowed from the symbol presentation to the choice Tasks specificities are indicated below each screen. See also **Appendix B.1** for a complete overview of all 6 experiments. **(B)** Experiment 1 payoff matrix. **(C)** Experiments 2-6 payoff matrix.

3.2.5 Matching probability and incentivization

In Experiment 2-6, participant's reports of confidence were incentivized via a matching probability procedure that is based on the Becker-DeGroot-Marshak (BDM) auction (Becker, Degroot, et al., 1964) Specifically, participants were asked to report as their confidence judgment their estimated probability (p) of having selected the symbol with the higher average value, (i.e. the symbol offering a 75% chance of gain (G75) in the gain conditions, and the symbol offering a 25% chance of loss (L25) in the loss conditions) on a scale between 50% and 100%. A random mechanism, which draws a number (r) in the interval [0.5 1], is then implemented to select whether the subject will be paid an additional bonus of 5 euros as follows: If $p \geq r$, the selection of the correct symbol will lead to a bonus payment; if $p < r$, a lottery will determine whether an additional bonus is won. This lottery offers a payout of 5 euros with probability r and 0 with probability 1-r. This procedure has been shown to incentivize participants to truthfully report their true confidence regardless of risk preferences (Hollard et al., 2016; Karni, 2009). Participants were trained on this lottery mechanism and informed that up to 15 euros could be won and added to their final payment via the MP mechanism applied on one randomly chosen trial at the end of each learning session (3×5 euros). Therefore, the MP mechanism screens were not displayed during the learning sessions.

3.2.6 Analysis

3.2.6.1 Variables

In all experiments, response time is defined as the time between the onset of the screen conveying the response mapping (Symbol for Exp. 1-2 & 6; Choice for Exp. 3-5; **Figure 3.1A**), and the key press by the participant. Confidence ratings in Exp. 1 were transformed form their original scale (0-10) to a probability scale, (50-100 %), using a simple linear mapping: confidence = (50 + 5 × rating)/100;

3.2.6.2 Statistics

All statistical analyses were performed using MatlabR2015a® (MathWorks). All reported p-values correspond to two- sided tests. T-tests refer to a one sample t-test when comparing experimental data to a reference value (e.g. chance: 0.5), and paired t-tests when comparing experimental data from different conditions.

Two-way repeated measures ANOVAs testing for the role of valence, information and their interaction were performed at the individual experiment level. One-way ANOVAs were used on main effects (e.g. individual averaged accuracy in gains minus losses) to test for the effect of experiments. Generalized linear mixed-effect (glme) models include a full subject-level random-effects structure (intercepts and slopes for all predictor variables). The models were estimated using Matlab's fitglme function, which maximize the maximum pseudo-likelihood of observed data under the model (Matlab's default option). Choice accuracy was modelled using a binomial response function distribution (logistic regression), whereas confidence judgments and response times were modelled using a Normal response function distribution (linear regression). For instance, the linear mixed-effect models for choice accuracy can be written in Wilkinson-Rogers notation as:

Choice_accuracy ~ 1 + Val. + Inf. + Val. * Inf. + Fix. + Stim. + Mask. + Sess. + (1 + Val. + Inf. + Val. * Inf. + Fix. + Stim. + Mask. + Sess. |Subject),

With Val: valence; Inf: information; Fix.: fixation duration (only available in Experiments 4-5); Stim.; stimulus display duration (only available in Experiment 5); Mask: Mask duration (only available in Experiments 4-5); Sess: session number.

Note that Val. and Inf. are coded as 0/1, but that the interaction term Val*Inf was computed with Val. and Inf. coded as -1/1 and then rescaled to 0/1. The robust regressions were performed with Matlab's robustfit function, using default settings. The algorithm uses iteratively reweighted least squares with the bisquare weighting function to decrease the impact of extreme data-points (outliers) on estimated regression coefficients.

3.3 Results

First, we evaluated the effects of our manipulation of the display and response settings across the experiments on average levels of choice accuracy and confidence ratings using multiple independent one-way ANOVAs. We found significant effects of the experiments on the average levels of accuracy ($F(5,102) = 5.72$, $p = 1.00\times10^{-4}$, $\eta^2 = 0.21$), mostly driven by a drop of accuracy in experiment 6 (see **Table 3.1** and **Appendix B.4: Figure B.4.1A**), but no effects on average levels of confidence ratings ($F(5,102) = 1.50$, $p = 0.1953$, $\eta^2 = 0.07$; **Table 3.1**). We also computed, at the session level (participants underwent 3 separate learning sessions per experiment), the correlations between confidence ratings and RT. When averaged at the individual level and tested at the population level (one sample t-test), this measure of the linear relationship between RT and confidence was very significant in all experiments (Exp. 1-6: all $ps < .01$; **Table 3.1**). The consistent negative and significant correlations across six experiments indicate that confidence is robustly associated with RT regardless of option-action mapping or time pressure manipulations, suggesting a strong link between instrumental and metacognitive processes. Yet, the correlation between confidence and RT was modulated by our experimental manipulations (effect of experiment: $F(5, 102) = 9.91$, $p < .001$, $\eta^2 = 0.32$) – post-hoc tests revealed that it was significantly altered by all our experimental manipulations in Exp. 3-6 (**Figure B.4.1B**).

Table 3.1. Demographics and behavior.

		Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5	Exp. 6
Gender	M/F	8/10	8/10	10/8	10/8	6/12	9/9
Age	mean ± STD	24.6 ± 8.50	24.6 ± 4.30	22.72 ± 3.24	23.84 ± 4.12	20.61 ± 1.77	22.35±3.49
Performance (accuracy; %)	mean ± SEM	76.50 ± 2.38	77.04 ± 1.69	80.00 ± 2.82	75.33 ± 2.34	73.40 ± 2.83	63.60 ± 2.88
Confidence (%)	mean ± SEM	79.19 ± 1.49	81.11 ± 1.58	78.78 ± 2.61	78.35 ± 2.24	78.09 ± 1.75	72.99 ± 2.14
Correlation (conf, RT)	mean ± SEM	-0.30 ±0.05	-0.41 ± 0.03	-0.18 ± 0.03	-0.16 ± 0.03	-0.10 ± 0.02	-0.12 ± 0.04
	t(17)	-5.31	-13.32	-5.55	-5.42	-4.87	-3.03
	(*p*-val)	(<.001)***	(<.001)***	(<.001)***	(<.001)***	(<.001)***	(0.008)**

The correlation between confidence and performance was performed at the session level using Pearson's R, then averaged at the individual level. Reported statistics correspond to a random-effects analysis (one sample t-test) performed at the population level.

STD: standard deviation. **SEM**: standard error of the mean. **T**: Student t-value.

* p <0.05; ** p <0.01; *** p <0.001

Next, we analyzed the effects of our experimental manipulation (valence and information) on the observed behavioral variables (choice accuracy, confidence, RT), using repeated measures ANOVAs in each individual study (**Figure 3.2**; **Table 3.2**). The parallel analyses of choice accuracy and confidence ratings replicated the results reported in (Fontanesi et al., 2019; Lebreton et al., 2019; Palminteri et al., 2015). Indeed, participants were more accurate in complete information contexts in five of six experiments (**Table 3.2**; main effect of information on accuracy, Exp. 1-5: ps < .05; Exp. 6: p = .1570). The effects of information on accuracy were actually not significantly different across our different experiments (**Appendix B.4: Figure B.4.2A**; effect of experiment: F(102,5) = 0.52, p = .7289, η^2 = 0.03). On the other hand, participants learned equally well in gain and loss contexts, as they exhibited similar levels of accuracy in gains and loss contexts in all experiments (**Table 3.2**; main effect of valence on accuracy, Exp. 1-6: all ps > .3; **Figure B.4.2A**; effect of experiment: F(5, 102) = 0.35, p = .884, η^2 = 0.02).

Despite similar performances in gain and loss contexts, and despite our attempt to cancel the valence-induced motor bias with our manipulations of the option-action mapping and time pressure, participants were slower in loss contexts in experiments 1-4 & 6 (**Table 3.2**; main effect of valence on RT: all ps < .01). These results not only replicate the results reported in (Fontanesi et al., 2019), but also assert the robustness of the valence-induced motor bias to the manipulation of response setups in human instrumental learning. Still, our experimental manipulations significantly reduced the motor bias in Exp. 3-5 (**Figure B.4.2C;** effect of experiment: F(5, 102) = 7.98, p < .001, η^2 = 0.28).

Table 3.2. Repeated measures ANOVA results reported separately for choice-relevant behavioral measures.

			Exp. 1	**Exp. 2**	**Exp. 3**	**Exp. 4**	**Exp. 5**	**Exp. 6**
Performance	**val.**	F(1,17), [η^2] (p-val.)	1.04, [0.01] (0.323)	0.00, [0.00] (0.971)	0.40, [0.00] (0.538)	0.01, [0.00] (0.912)	0.33, [0.00] (0.571)	0.37, [0.04] (0.553)
	inf.	F(1,17), [η^2] (p-val.)	4.28, [0.04] (0.054)~	18.64, [0.15] (0.001)***	5.56, [0.04] (0.031)*	3.26, [0.06] (0.089)~	10.17, [0.07] (0.005)**	2.19, [0.02] (0.157)
	val.×inf.	F(1,17), [η^2] (p-val.)	1.06, [0.01] (0.319)	0.77, [0.01] (0.393)	0.06, [0.00] (0.816)	4.36, [0.04] (0.052)~	1.04, [0.01] (0.326)	3.57, [0.02] (0.075)~
Confidence	**val.**	F(1,17), [η^2] (p-val.)	33.11, [0.27] (<.001)***	15.43, [0.19] (0.001)**	12.18, [0.03] (0.003)**	19.14, [0.07] (<.001)***	16.71, [0.15] (<.001)***	26.71, [0.12] (<0.001)***
	inf.	F(1,17), [η^2] (p-val.)	2.00, [0.00] (0.175)	4.92, [0.02] (0.040)*	2.28, [0.02] (0.149)	3.21, [0.01] (0.091)~	11.07, [0.01] (0.004)**	0.11, [0.00] (0.743)
	val.×inf.	F(1,17), [η^2] (p-val.)	7.58, [0.02] (0.014)*	4.25, [0.01] (0.055)~	1.61, [0.01] (0.222)	4.46, [0.01] (0.050)~	7.87, [0.02] (0.012)*	5.16, [0.01] (0.036)*
RT	**val.**	F(1,17), [η^2] (p-val.)	13.25, [0.03] (0.002)**	13.15, [0.08] (0.002)**	12.47, [0.01] (0.003)**	11.23, [0.01] (0.004)**	1.97, [0.00] (0.178)	15.56, [0.02] (0.001)**
	inf.	F(1,17), [η^2] (p-val.)	0.12, [0.00] (0.733)	7.64, [0.01] (0.013)*	1.82, [0.00] (0.195)	0.31, [0.00] (0.586)	0.09, [0.00] (0.766)	3.60, [0.00] (0.074)~
	val.×inf.	F(1,17), [η^2] (p-val.)	4.94, [0.01] (0.040)*	0.36, [0.00] (0.558)	1.32, [0.00] (0.266)	2.32, [0.00] (0.146)	0.70, [0.00] (0.414)	2.02, [0.00] (0.173)

val.: valence; **inf.**: information;

~ p <0.1; * p <0.05; ** p <0.01; *** p <0.001

Importantly, despite similar performance in gain and loss contexts, participants were less confident in loss contexts (**Table 3.2**; main effect of valence on confidence, Exp. 1-6: all *ps* < .01), with very similar effect sizes across all experiments (**Figure B.4.2B**; $F(5, 102) = 1.26$, $p = .289$, $\eta^2 = 0.06$). These effects were mitigated when more information was available (**Table 3.2**; interaction valence × information on confidence: all *ps* < .05). These results not only replicate those reported in Lebreton et al., 2019, but also assert the robustness of the valence-induced confidence bias.

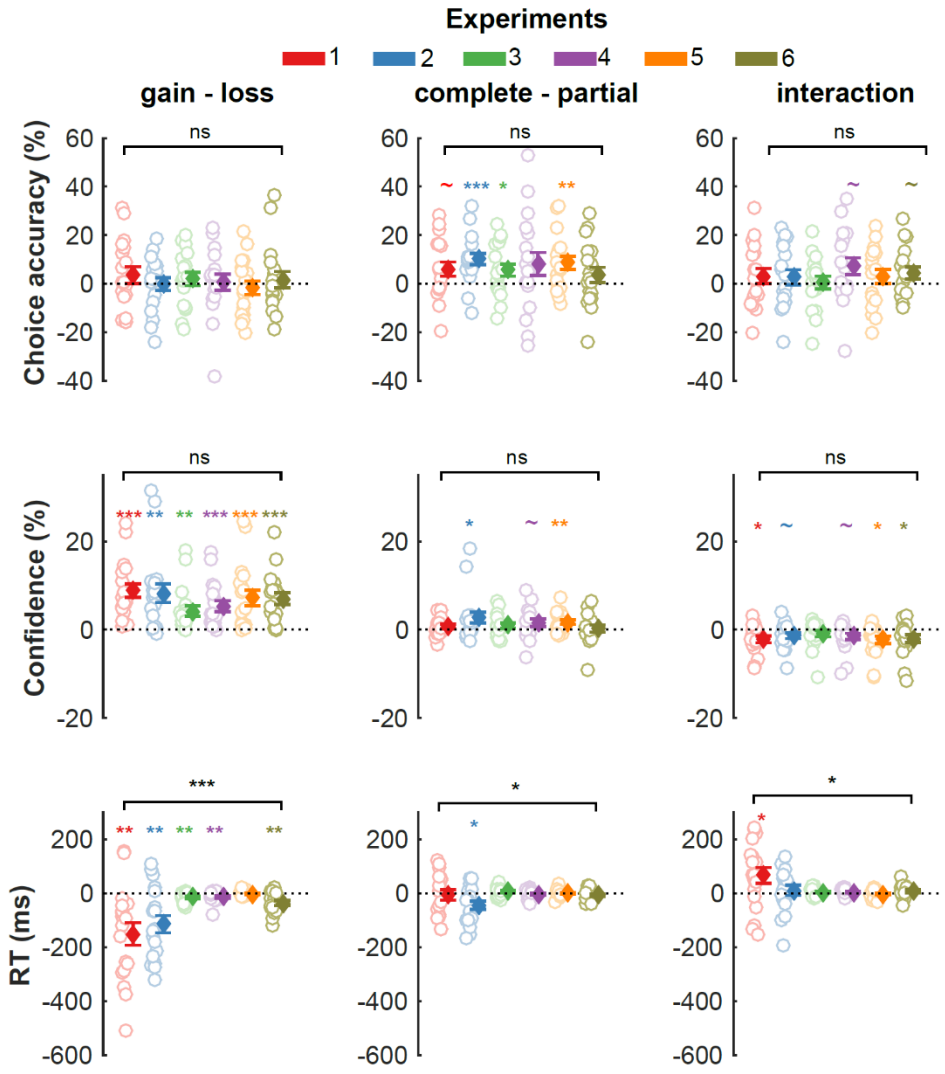


Figure 3.2. Behavioral results. Effects of the main manipulations (left: valence; middle: information; right: interaction) on relevant measures of choice-relevant behavior (top: performance; middle: confidence; bottom: response times). Analyses are independently performed in the six different experiments using repeated-measures ANOVAs. Empty dots with colored edges represent individual data points across different experiments; filled diamonds and error-bars represent sample mean ± SEM. The horizontal bar indicates a one-way ANOVA testing the effect of experiment on each manipulation (see **Appendix B.4** for details).
~ p <0.1; * p <0.05; ** p <0.01; *** p <0.001

Overall, the analyses of the data collected in six different versions of our experiment (N = 108) clearly underline the remarkable robustness of the effects of outcome valence on both confidence and RT. Only one experimental condition succeeded in cancelling the valence-induced motor bias (Experiment 5). Note that in this experiment, we still observed the confidence bias as evidenced by a significant main effect of valence on confidence (**Table 3.2**; $F(1,17) = 16.71$, $p < .001$, $\eta^2 = 0.15$), but not on RT ($F(1,17) = 1.97$, $p = .178$, $\eta^2 = 0.001$). This suggests that the effects of outcome valence on confidence and RT are – partly – dissociable. In other words, we can observe a lower confidence in loss contexts, even when RTs are indistinguishable from gain contexts.

In order to give a comprehensive overview of the relationship between accuracy, confidence and RT, and to quantify the effects of the different available predictors on these behavioral measures, we also ran generalized linear mixed-effect regressions. Independent variables included not only valence, information and their interaction, but also the different available timings (e.g. duration of the stimulus or mask display) and a linear trend accounting for the session effects (see methods for details). These sensitive trial-by-trial analyses replicated the main ANOVA results reported above regarding the effects of valence and information on performance, confidence and RT (**Figure 3.3**; **Appendix B.3**). They also confirmed that, in experiment 5, no effect of valence can be detected on RT and performance ($p = .349$ and $p = .620$) while a robust effect is observed on confidence ($p = .002$).

We also ran an additional mixed model, which estimated the effect of our experimental factors on confidence, while controlling for RTs – i.e. including RTs in the dependent variables (**Table B.3.4**). Importantly, and replicating previous findings (Lebreton et al., 2019), the main effect of valence on confidence remained significant in all experiments ($p < .001$), providing additional evidence that the valence-induced confidence bias is partially dissociable from the valence-induced motor bias.

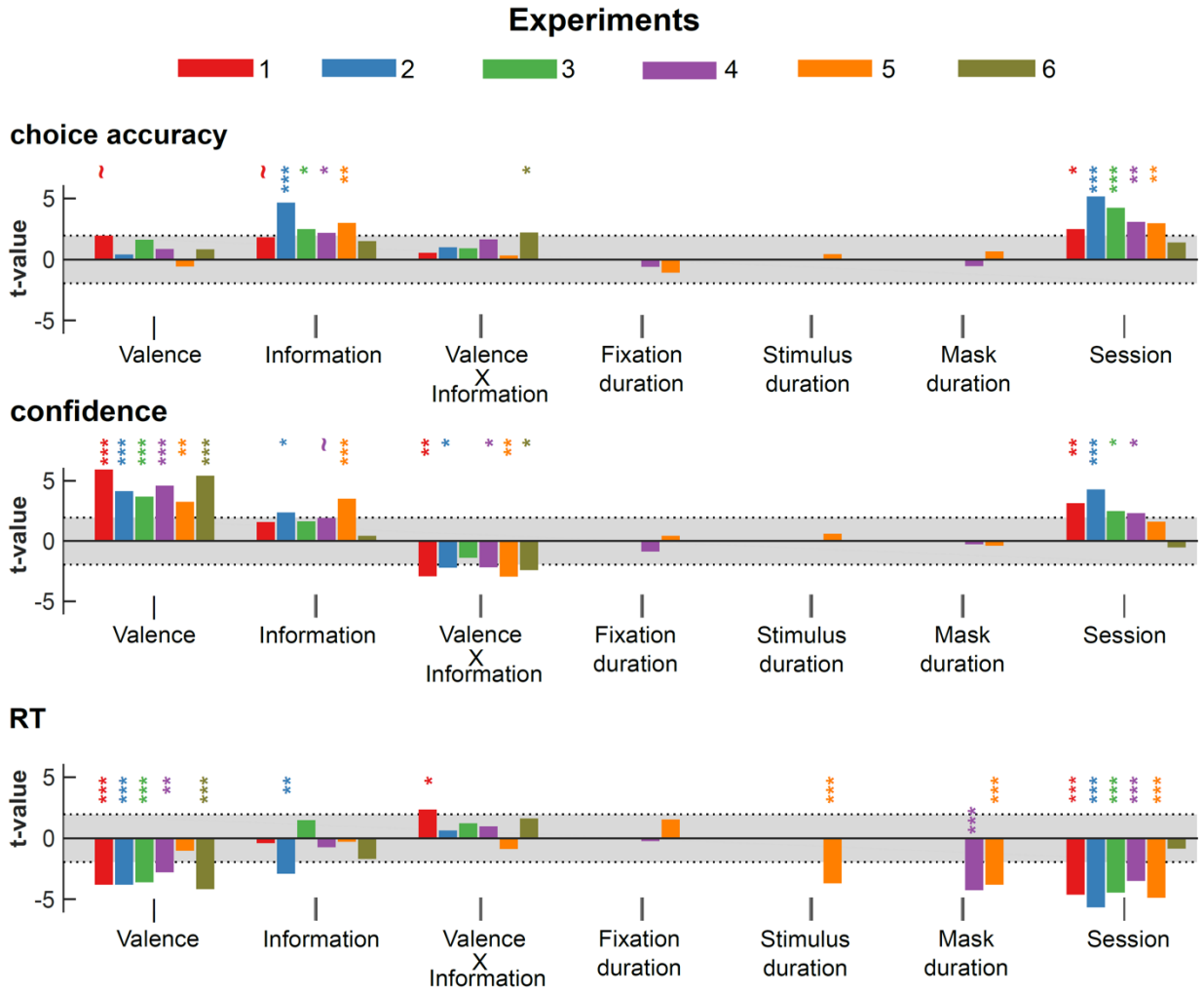


Figure 3.3. Generalized linear mixed-effects models. Estimated standardized regression coefficients (t-values) from generalized linear mixed-effects (GLME) models, fitted in the different experiments. Top: logistic GLME with performance as the dependent variable. Middle: linear GLME with confidence as the dependent variable. Bottom: linear GLME with RT as the dependent variable; Shaded area represent area where coefficients are not significantly different from 0 (abs(t-value)< 1.95; p >.05).
~ p<.10; * p <.05; ** p <.01; *** p <.001

Because the valence-induced motor bias – i.e. the slowing down of RTs in loss compared to gain contexts – was extremely robust to our experimental manipulations aiming at cancelling it, the ANOVA and regressions above provide only limited evidence on whether valence-induced decreasing on confidence can be observed in the absence of the valence-induced slowing of RT. In the following paragraphs, we therefore used a different analytical strategy leveraging inter-individual differences to test this hypothesis. We assessed the link between individual slowing down (RT in gain – loss) and individual decreases in confidence (confidence in gain-loss) in our full sample and in each individual study using robust linear regressions (see methods for details). In those regressions, the coefficients for the intercept and slope quantify two different but equally important signals: First, the y-intercept represents a theoretical individual who exhibits no effect

of valence on RT (RT in gain – loss = 0, **Figure 3.4A**): an intercept significantly different from 0 therefore indicates that a significant effect of valence on confidence can be observed in the absence of an effect on RT. Second, the slope quantifies how the effect of valence on confidence linearly depends on the valence-induced slowing of RT. Both at the population level (i.e., combining data from all six experiments) and in each individual study, the intercepts of those regressions were estimated to be significantly positive (all $ps < .05$; **Figure 3.4 A-B; Appendix B.5**). This indicates that valence-induced changes on confidence are detectable when valence induced-changes on RT are absent. Note that at the population level, the slope of the regression was also significantly negative ($\beta = -0.02 \pm 0.01$, $t(106) = -3.75$, $p < 0.001$), indicating that, compared with the gain context, the more participants were slowed down by the loss context, the less confident they were in their response. Therefore, the valence-induced motor and confidence biases are only partially dissociable.

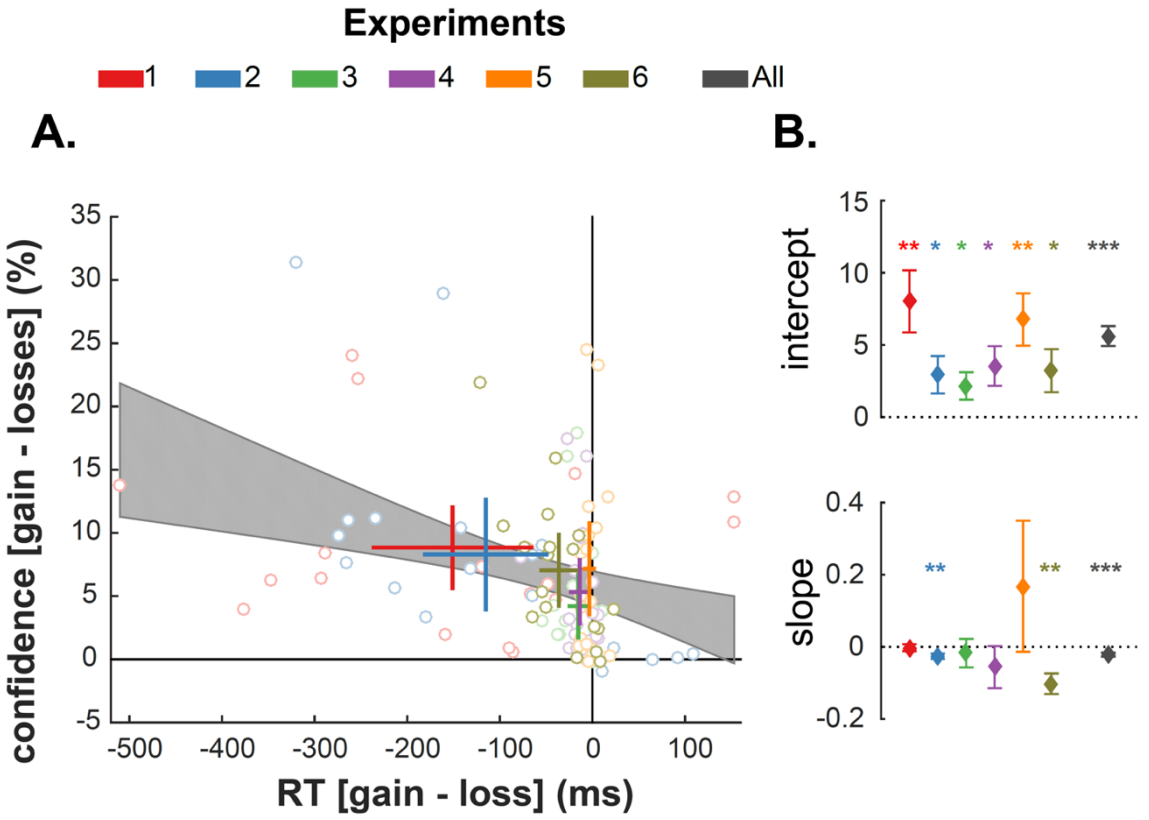


Figure 3.4. Assessing the link between the effects of valence on confidence and response times. **(A)** Inter-individual correlations between the effects of valence on confidence (Y-axis) and response times (X-axis) across experiments. Dots represent data points from individual participants. Thick lines represent the mean ± 95%CI of the effects of valence on confidence (vertical lines) and response times (horizontal lines). Experiments are indicated by the dot edge and line color. The black shaded area represents the 95%CI of the inter-individual linear regression. Note that potential outliers did not bias the regression, given that simple and robust regressions gave very similar results. **(B)** Results from inter-individual regressions of the valence-induced RT slowing on the valence-induced confidence difference across different experiments. Top: estimated intercepts of the regressions. Bottom: estimated slopes of the regressions. Diamonds and error-bars represent the estimated regression coefficients (β) and their standard error.

* p<.05; ** p <.01; *** p <.001

3.4 Discussion

The present work investigated the relationship between valence-induced biases affecting two different behavioral outputs: response time and confidence. We confirm here, in 6 variations of a simple probabilistic reinforcement-learning task, that learning to avoid punishment increased participants' response time (RT) and decreased their confidence in their choices, without affecting their actual performance (Fontanesi et al., 2019 ; Lebreton et al., 2019). The valence-induced bias on RT is currently interpreted as a manifestation of a motor – or instrumental - Pavlovian bias (Boureau & Dayan, 2011; Guitart-Masip et al., 2012). In the associative learning literature, similar Pavlovian effects - whereby the presentation of reward-associated stimuli can motivate behaviors that have produced rewards in the past – have been described (Mahlberg et al., 2019). One of the most studied effect is the Pavolvian-Instrumental Transfer (PIT), which is defined as an increased vigor in instrumentally trained responses when these are made in the context of Pavlovian, or reward-associated, cues (Cartoni et al., 2016; Holmes et al., 2010). While we did not employ standard PIT procedures in the current studies, which would involve separate Pavlovian and transfer phases (Colwill & Rescorla, 1988; Rescorla & Solomon, 1967; Watson et al., 2014), our findings nonetheless parallel those from Pavlovian-Instrumental Transfer studies, by showing faster reaction times in the context of reward, but not punishment cues.

The valence-induced decrease in confidence has been described as a value-to-confidence contamination, potentially generated by a mechanisms of affect-as-information (Lebreton et al., 2018; Schwarz and Clore, 1983). Note that some authors have warned about possible mis-identifications between a true confidence bias and a change in metacognitive sensitivity (Fleming & Lau, 2014). Yet, because we previously established in a perceptual task that the outcome valence manipulation specifically impacts the confidence bias and not metacognitive sensitivity (Lebreton et al., 2018), we assume that that the same experimental manipulation produces similar effects in a reinforcement-learning task.

One of the motivations behind the present chapter was to rule out a potential alternative explanation of the observed decrease in confidence: participants could derive confidence estimates by monitoring changes in their own response times. Indeed, because it has been suggested that humans can infer confidence levels from observing their RT (Desender, Opstal, et al., 2017; Kiani et al., 2014), the valence-induced bias on confidence could be spuriously driven by a valence-induced motor bias operating at the level of motor initiation (Boureau & Dayan, 2011; Guitart-Masip et al., 2012). As such, valence-induced confidence biases would then merely reflect a secondary effect of valence mediated by response time slowing, and not a primary meta-cognitive bias. Crucially, this possibility is not ruled out by previous studies, where effects of affective states on confidence judgments in perceptual or cognitive tasks typically lacked control over RT (Giardini et al., 2008; Koellinger and Treffers, 2015; Massoni, 2014, but see Lebreton et al., 2018). We address this issue in the current set of experiments by dissociating decisions from motor mapping, thereby partially removing the association between RT and confidence.

We analyzed six datasets composed of two published datasets (Exp. 1-2) and four new experiments (Exp. 3-6). Over those six experimental datasets, the first noticeable result is that we systematically replicated previous instrumental learning results using the same paradigm with very consistent

effect sizes (Palminteri et al., 2015, 2016): participants learn equally well to seek reward and avoid punishment, and learning performance benefits from complete information (i.e. feedback about the counterfactual outcome). The reliability of the results extended beyond choice behavior as confidence and RT were, respectively, lower and slower in punishment contexts compared to reward contexts, as previously reported (Fontanesi et al. 2019; Lebreton et al., 2019), thus confirming the robustness of the valence bias.

The second important result is that the slowing down of RTs in loss contexts is extremely resilient, as it was still observed when the mapping between motor response and option selection was dissociated by our experimental design (Exp. 3-4) and when significant time pressure was applied on the decision (Exp. 6) – albeit with significantly lower effect sizes. This result speaks to the strength and the pervasiveness of the valence-induced bias operating at the motor level (Boureau & Dayan, 2011; Guitart-Masip et al., 2012).

Third, and importantly, we still observed a significant valence effect on confidence when the valence effects on RT were dramatically reduced (Exp. 3, 4 and 6) or absent (Exp. 5), indicating that the lower confidence observed in the loss-avoidance context is – at least partly– dissociable from the concomitant slowing down of motor responses. This was confirmed by additional evidence from inter-individual difference analyses, showing that in all six experiments, a theoretical subject exhibiting no valence-induced bias in RT would still exhibit a valence-induced bias in confidence. Note that the absence of a significant motor bias observed in Exp. 5 could be caused by the successful changes in the experimental setup, that were implemented with this specific goal in mind. Yet, it could also be a false negative: the experimental setup could still be inefficient to cancel the motor bias, but the sampled participants just happened – by chance – to not exhibit the motor bias. Regardless of the reason for this null-effect, the important point is that in this sample – where we failed to detect a significant effect of valence on reaction times – there was still an effect of valence on confidence. Altogether, these results suggests that it is unlikely that the valence-induced bias on confidence reported in human reinforcement-learning (Lebreton et al., 2019) is a mere consequence of a response time slowing caused by an aversive motor Pavlovian bias. Our results are also consistent with recent findings (Dotan et al., 2018b) challenging the notion that humans infer confidence levels purely from observing their own response times, and suggesting that decision reaction times are a consequence rather than a cause of the feeling of confidence (Desender, Opstal, et al., 2017; Kiani et al., 2014). It is worth noting that in most studies, decision-time (i.e. when participants reach a decision) and response times (when participants indicate their choice) are not experimentally dissociated and often conflated in the same measure. Here we delayed the mapping between decisions (in the option space) and action selection (motor space), which resulted in an effective control over response times. Future studies will investigate whether participants can keep track of an internal measure of decision time, which could influence confidence. Likewise, we cannot pretend that our experimental manipulations removed all valence (Pavlovian) effects on motor responses. We only managed to modulate one component of our participants' response vigor: the response times (RT).

In our data, we also observed that confidence ratings and RT are robustly associated regardless of time pressure manipulation. The negative correlation between confidence and RT was consistently found in over six experiments. This coupling is consistent with predictions from most sequential-

sampling models (van den Berg et al., 2016; De Martino et al., 2013; Navajas et al., 2016; Pleskac and Busemeyer, 2007; Ratcliff and Starns, 2009, 2013; Yu et al., 2015), which posit that confidence and RT jointly emerge from a single mechanism of evidence accumulation. Importantly, we still observed robust correlations between confidence and motor RTs when we dissociated action selection from the option evaluation. Therefore, the motor execution of a decision might be more important than previously thought in sequential-sampling models of confidence, which mostly focus on decision times.

The replicability and robustness of the valence-induced confidence bias implies that the manipulations of valence could prove useful to dissociate fundamental components of decision-making and metacognitive judgment, such as objective uncertainty and subjective confidence (Bang & Fleming, 2018). The dissociation between objective uncertainty and subjective confidence is anticipated by post-decisional and second-order models of confidence (Fleming & Daw, 2017; Pleskac & Busemeyer, 2007), which postulate that confidence is formed after the decision and thereby might be influenced by other internal or external variables (Moran et al., 2015; Navajas et al., 2016; Yu et al., 2015). It is worth noting that our results do not rule out the possibility that RT is used to guide metacognitive judgment of confidence before and after the decision. Actually, the fact that participants who exhibit the strongest valence-induced motor bias are also the ones that exhibit the strongest confidence bias (significant negative slope(s) in **Figure 3.4 A-B** and **Appendix B.5**) indicates that their reaction times and confidence are linked. Observing one's RTs could therefore be one of the factors that influences confidence after the decision was made, as posited in second-order models.

In a previous study (Fontanesi et al., 2019), we analyzed the effects of valence on RT, on a different dataset collected with a similar experimental design – although omitting confidence judgments. There, using an approach combining reinforcement-learning and decision-diffusion modelling, we reported that valence influences two critical parameters of the response time model: the non-decision-time - which typically represents perceptual and motor processes – and the decision threshold – which indexes response cautiousness. We speculate that this distinction is relevant to interpret the results of the present report. We propose that the portion of the valence-induced response time slowing that we were able to cancel through response-mapping manipulation could be linked to the non-decision-time modulation; on the other hand, the residual irreducible valence-induced response time slowing could be linked to the increased response cautiousness. Yet, given the disruption of the response mapping present in most experiments in the current study, the combined reinforcement-learning and decision-diffusion modelling approach cannot be applied to the present data to test this hypothesis. Further experiments are therefore needed to refine the computational description of valence-induced biases in reinforcement-learning, and their consequences on performance, confidence and response times.

Finally, the question arises to what extent incentive-related, confidence, and Pavlovian and instrumental processes, which all influence behavior in the current study, are supported by dissociable, or overlapping brain systems. Incentives are typically processed by the brain reward system, of which the ventral striatum (VS) and ventromedial prefrontal cortex (vmPFC) are key structures (Bartra et al., 2013; Haber & Knutson, 2009; Pessoa & Engelmann, 2010). The anterior insula is also often involved in incentive processing, and seems to preferentially code negative

incentive value (Bartra et al., 2013; Engelmann et al., 2015, 2017; Palminteri, Justo, Jauffret, Pavlicek, Dauta, Delmaire, Czernecki, Karachi, Capelle, & Durr, 2012). This set of neural structures is also involved in the computation of positive (vmPFC, VS) and negative (anterior insula) reward prediction errors (RPE)s. RPEs are an essential part of reinforcement learning models of Pavlovian and instrumental learning, and reflect the difference in expected and observed rewards (or punishments), which is used to update future decision value estimates. Unsurprisingly, brain regions associated with Pavlovian Instrumental Transfer also involve these regions associated with processing predominantly appetitive stimuli, i.e. the ventral striatum and ventral region of the prefrontal cortex, but also regions associated with predominantly aversive stimuli, i.e. the amygdala (Cartoni et al., 2016; Holmes et al., 2010; Deborah Talmi et al., 2008). Interestingly, recent neuroimaging studies have also shown that neural signals in the vmPFC correlate with confidence judgments in a variety of tasks (De Martino et al., 2013; Lebreton et al., 2015; Shapiro & Grafton, 2020). Taken together, there is significant overlap in the neural systems that support incentive processing (VS, vmPFC) and appetitive Pavlovian and instrumental learning (VS), on the one hand, and confidence (vmPFC) on the other. Note further, that ventral striatum is situated in the basal ganglia and has direct projections with vmPFC (Haber & Knutson, 2009), and can therefore function as an interface between motor and affective/motivational systems. Regions encoding incentives and learning in the aversive domain do not seem to share the same direct interconnectivity with vmPFC and motor regions (Cerliani et al., 2012). The concurrent representation of key cognitive processes in sub-regions of the reward system, together with its connectivity profile, make it a good candidate to explain the valence-induced motor and confidence biases observed in the current study. Note, however, that these are merely neuroanatomical hypotheses based on integrating results from related literatures on reward, reinforcement learning and PIT. It is therefore essential that future neuroimaging research identifies the underlying neurobiological basis of the valence-induced motor and confidence biases we demonstrate here.

CHAPTER 4

The Neural Mechanisms of Confidence in Reward Learning and Punishment Learning

4.1 Introduction

Previous studies (Lebreton et al., 2018, 2019) and our experiments (Chapter3; Ting et al., 2020b) have consistently demonstrated the effect of outcome valence (i.e., gain and loss) on confidence judgment in reinforcement learning and other tasks. Specifically, confidence about a choice being correct (i.e., choosing the option with higher expected value) is higher when learning to gain rewards compared to learning to avoid losses, even though learning performance is equal across conditions. While this finding is robust across studies and contexts, the neural correlates of this valence-based confidence bias have not been addressed to date. To answer this question, the present chapter aims to investigate whether the same cognitive and neurological mechanisms underlying confidence formation and confidence bias are required for reward and punishment learning.

Reinforcement learning is a dynamic process of value updating integrating many sources of information, such as uncertainty in the environment, motivation and performance history. These sources of information are usually associated with confidence judgments, which is defined as belief- or subjective probability- of choosing the better option given the available evidence (Meyniel, et al., 2015; Pouget et al., 2016; Sanders, et al., 2016). Three candidate neural networks are involved in confidence formation in learning. The first neural network consists of Insula, dorsal anterior cingulate cortex (dACC) and inferior frontal gyrus (IFG). This network was linked to environmental uncertainty and task difficulty, but also negatively correlated with confidence (Fleming et al., 2012, 2018; Hebart et al., 2016; Hilgenstock et al., 2014; Shenhav et al., 2014). The second candidate is the neural network processing incentives and subjective values, such as ventral striatum (VS) and medial prefrontal cortex (mPFC). Recent studies have strongly supported the role of these regions on encoding neural confidence signals (Bang & Fleming, 2018; De Martino et al., 2013, 2017; Fleming et al., 2018; Hebart et al., 2016; Lebreton et al., 2015). The third potential regions of interest are perigenual anterior cingulate cortex (pgACC) and medial temporal lobe (MTL), which have been associated with self-performance evaluation and memory-retrieval, respectively (Chua et al., 2014; Wittmann et al., 2016). These regions have been shown to positively encode confidence ratings (Bang & Fleming, 2018; Chua et al., 2009; Kim & Cabeza, 2009). However, these confidence-related neural networks were established on natural or positive environments and consequently, there is a lack of clarity about the neural mechanisms underlying confidence formation in aversive environments. According to the current understanding of neural confidence signals, two specific research questions were formed: (1) which regions encode general confidence signals (i.e. irrespective of context), and (2) whether the confidence encoding is context-dependent?

Previous studies have shown two main features of confidence processing in the brain. Firstly, confidence is automatically processed and integrated to value system, such as ventral medial prefrontal cortex (vmPFC) (De Martino et al., 2013, 2017; Lebreton et al., 2015; Lopez-Persem et al., 2020; Shapiro & Grafton, 2020). Specifically, the correlation between vmPFC activity and confidence can be found even when stating confidence is not explicitly required (Lebreton et al., 2015; Lopez-Persem et al., 2020; Shapiro & Grafton, 2020). This finding suggests that confidence judgments are formed and encoded in the brain before explicitly making a choice and reporting confidence. Secondly, the confidence judgment represents the subjective feeling and is generally

processed in the valuation system. The feature is in line with the common currency theory (Montague and Berns, 2002; Levy & Glimcher, 2012) which suggests that value of different classes of goods (e.g., attraction of faces, probability of future event, confidence about making better choice) are encoded and integrated under a common scale in the same network. Although this feature suggests that confidence ratings are processed in the common neural network regardless of experimental manipulations, the overlap in neural systems implies that confidence might interact with ongoing value signals, such as option value and context value.

Recent studies have shown that the option value and context value are updated simultaneously in the reinforcement learning (Palminteri et al., 2015). Moreover, context value serves as a reference point to adapt feeling to the received outcome (Palminteri et al., 2015). The context value also carries affective information (e.g., positive and negative), which is likely to interact with cognitive and metacognitive functions. This idea is supported by the brain imaging studies showing affective information modulates the activity in the regions associated with metacognition (Fleming et al., 2012, 2014; Lebreton et al., 2015) and emotion-cognition integrations (Pessoa, 2008). For example, PFC and VS are intimately involved in computing subjective value (Bartra et al., 2013; Levy & Glimcher, 2012; Perlstein et al., 2002) and Insula is involved in processing undesired events (Engelmann et al., 2015; Tom et al., 2007). The examples place these regions in a strategic position to integrate affective information and valuation processes with metacognition. In addition, Clore & Huntsinger, (2007) also suggests affective information might influence metacognition as negative affect facilitates deliberative thinking and might delay confidence formation and improve metacognitive functions.

The neural circuitry underlying confidence formation should be investigated with some caution because confidence is usually associated with a variety of ongoing value signals. To solve this issue and to investigate how monetary-driven affective context is involved in confidence formation, one plausible method is to distinguish effects of outcome valence through experimental design and then investigate neural confidence signals under different valence-driven affective contexts (i.e., pure gain and pure loss contexts). Therefore, we adopted the optimized learning task from Chapter 3 (Ting et al., 2020b),which dissociated effect of outcome valence and measurements (i.e., learning performance, reaction time and confidence) for use in the MRI scanner. In order to find the neural mechanisms of confidence formation under different affective contexts, we applied both computational modeling and neural imaging results to test if the context value was processed. Afterward, we tested where and how confidence was processed in the brain. We recruited 40 participants and successfully replicated behavioral results that we demonstrated in Chapter 3 and recent studies (Fontanesi et al., 2019; Lebreton et al., 2019; Ting et al., 2020b). The neural imaging results showed brain value system was generally involved in confidence processing before participants explicitly stated confidence. The further ROI analysis indicated that confidence ratings from loss contexts were not consistently encoded compared to the confidence ratings in gain contexts.

4.2 Material and Methods

4.2.1 Participants

40 participants (female = 23; Age = 22.69±4.44) were recruited through official website and poster adverts distributed on campus. The ethic approval was obtained from the Psychology department at the University of Amsterdam (reference number: 2018-EXT-9205). Before the experiment, we used a prescreening procedure and only participants that passed this were invited to come to the MRI scanner and were sent an invitation email and detailed information about the experiment and MRI. Participants were asked to arrive laboratory 30-min before the experiment. Once participants arrived, they gave informed consent and read instruction again. Afterward, they experienced a 16-trial practice with the same learning task (but using different symbols) as well as lottery incentivize procedure (see below).

4.2.2 Probabilistic instrumental tasks

We adopted our previous instrumental reinforcement learning task (Lebreton et al., 2019; Palminteri et al., 2015; Ting et al., 2020b) for fMRI and by adding incentivized confidence ratings in each trial. During the learning task, participants were asked to maximize payoff by choosing the symbol with the higher expected value in a pair at each trial (**Figure 4.1A**). Totally four fixed pairs of abstract symbols were used to represent four conditions in the two (feedback valence: gain or loss) by two (information: partial or complete) within-subjects design (**Figure 4.1B**). Specifically, eight symbols were divided into four fixed combinations and are constantly arranged to gain/partial (Gp), loss/partial (Lp), gain/complete (Gc) and loss/complete (Lc) conditions. Each pair of symbols not only indicated a specific condition but also possible outcomes. For example, for gain contexts (i.e., Gp and Gc), the possible outcomes are +€1 or +€0.1. In contrary, -€1 or -€0.1 are possible outcomes in the loss contexts (i.e., Lp and Lc). The probabilistic outcome of an option was determined by reciprocal probabilities, 75% or 25% (**Figure 4.1C**). The symbol that enjoys higher expected value ($\sum \text{probability} \times \text{outcome}$) was defined as correct option in each pair. Note that only the chosen outcome was added to the final payoff in both the incomplete and complete feedback conditions.

All the participants completed three runs of 80 trials as each condition (i.e., each pair of symbols) repeated 20 times (**Figure 4.1A**). In each trial, the symbols were presented first (1500-3500ms; mean = 2050ms). To avoid the potentially confounding influence of motor responses during symbol evaluation, the symbols disappeared for a while (500-3000ms; mean = 800ms) after symbol presentation. Afterwards, two white bars appeared on either right or left of the symbol to indicate which button should be pressed to select corresponding symbol (i.e., right button: white bar was on the right side of symbol), even though the symbols were invisible. Once decision was made, two red bars were displayed besides the chosen symbol (500ms). Before seeing the outcome, participants were asked to state their confidence about choosing the symbol that is better on average. Confidence ratings were done on a scale ranging from 50% to 100% with incremental steps of 5%, and without time constraints. At the end of each trial, participants were shown the outcome from the chosen option only in the partial information conditions (i.e., Gp and Lp) for 2000ms. Otherwise, both chosen and unchosen outcome were displayed in the complete information conditions (i.e., Gc and Lc).

In order to motivate participants to accurately report confidence, confidence judgments were incentivized by a Matching Probabilities (MP) mechanism, a well-validated method from behavioral economics adapted from the Becker-DeGroot-Marschak auction (Becker, DeGroot, et al., 1964; Ducharme & Donnell, 1973b). Specifically, we randomly selected three trials from three runs (i.e., one trial/ run) and then compared the confidence rating p at that trial with a random number r (chosen from the range between 50% and 100%). If $p \geq r$, then participants won the bonus of 5€ when the chosen symbol indeed had the higher expected value (i.e., the correct one), otherwise, participants won nothing. If $p < r$, participants won the bonus of 5€ with probability of r, otherwise, won nothing with probability of 1-r. The euros earned from game was exchanged to the actual money with a certain exchange rate (1 EU in game = 0.3 payout EU). Again, all participants were clearly informed about the rule of payout and experienced practice trials in both the learning task and confidence incentivization before the real experiment in the MRI scanner.

After the learning task, participants were instructed to perform an additional task outside of the scanner. In the post-learning transfer task, each symbol from the last run was paired with all other 7 symbols, thus forming 28 new pairs. Participants were asked to choose one symbol that can benefit them more. Unlike the learning task in the main experiment, no feedback and monetary incentives were offered in this task. However, participants were asked to imagine that they were able to earn money from the chosen symbols. The final payout was computed as follows: show-up fee (20€), accumulated outcome from the learning task and bonus from confidence incentivization procedure. The mean and standard deviation of payout were 32.18±3.46€. All the tasks were implemented using MatlabR2015a® (MathWorks) and the COGENT toolbox.

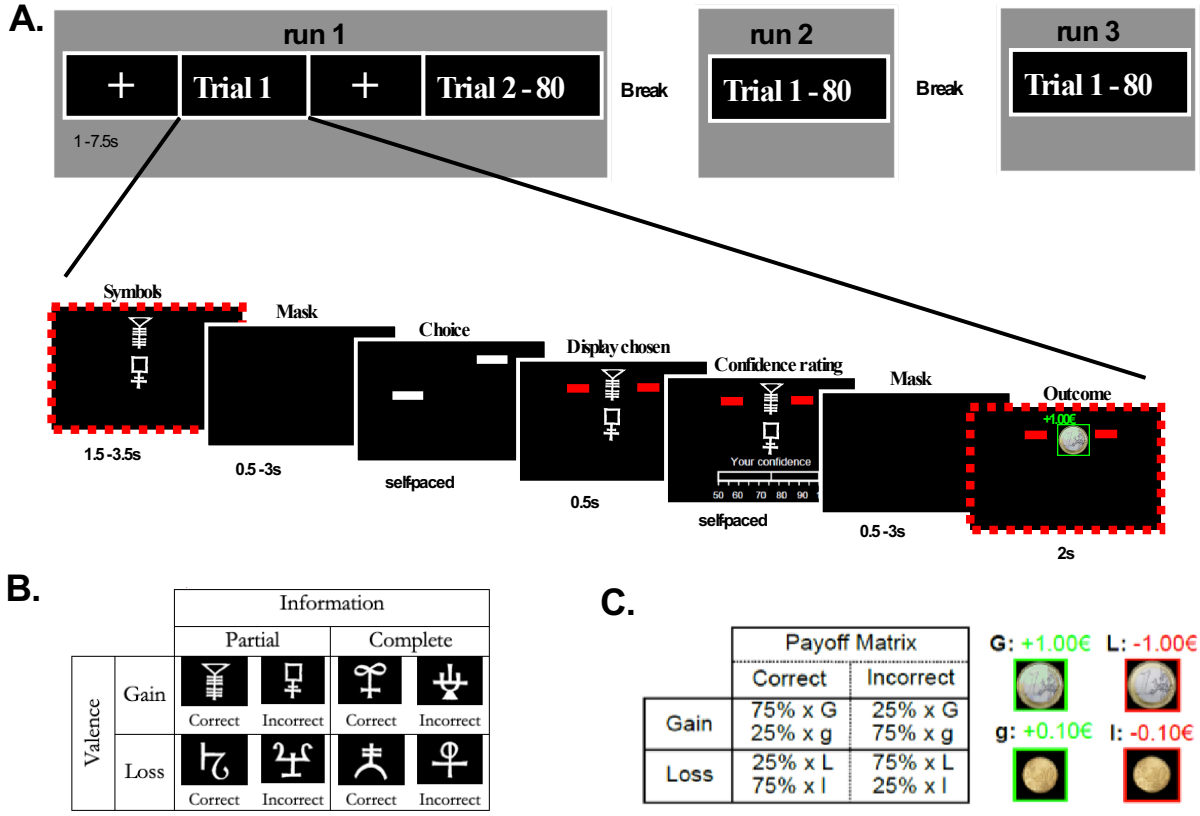


Figure 4.1. Experimental design. (A) Upper: Three runs in the scanner and example of gain partial trial. Four conditions were presented 20 times in a pseudorandom manner and consequently formed total 80 trials for each run. Two trials were separated by a cross-fixation with a random inter-trial interval selected from the range between 1s and 7.5s. After each run, subjects were asked to take a break for maximum 1 minute in the scanner and stay still. Lower: the example trial in the gain/partial condition. The red dashed rectangle represents the period of interest in the fMRI analyses. **(B)** Illustration of two-by-two factorial design with outcome valance and information manipulations. Each condition was consistently associated with a pair of symbols in each run. One symbol was the correct option as it had a relatively higher expected value. The set of symbols were not the same for each run, so that subjects have no prior experiences about symbols at the beginning of each run and have to re-learn associations throughout the run. All the symbols were taken from Agathodaimon alphabet. **(C)** The payoff matrix showed probability of getting larger gains and smaller gain in the gain conditions, and probability of getting larger loss and smaller loss in the loss conditions.
G: larger gain; g: smaller gain; L: larger loss; l: smaller loss

4.2.3 fMRI data acquisition

The fMRI data was acquired using 3.0-Tesla Philip Achieva scanner with 32 channels head array coil. We recorded both structural images and functional brain images. T1 weighted structural scans were recorded with following parameters: FOV (Field of View): 240×180×220 mm^3, Voxel size =1×1×1 mm^3, TR = 8.2ms and TE = 3.7ms. Each T2*-weighted functional scan consisted of 36 axial echo-planar images (EPI) acquired in ascending sequence with voxel size of 3×3×3 mm^3, slice gap = 0.3 *mm*, *T*R= 2000ms, TE = 28ms and the flip angle of 76°. Each subject completed 3 runs in a scanning session. Given the task was self-paced and fMRI scanner was manually terminated (i.e., ~10 second after the last feedback phase), the total numbers of functional scans for each subject in each run were not the same. Most participants completed the task around 15 minutes. The field maps (i.e., magnetic field's inhomogeneity) was collected as well between second and third run.

4.2.4 Behavioral analyses

In this study, the dependent variables of interest were learning performance, reaction times, confidence ratings and their relationships. The learning performance referred to the probability of choosing the relatively better symbol (i.e., the one with higher expected value) in each pair of symbols. These measures were averaged over three runs for each condition. Aggregated data were calculated for each condition and participants and were then fed into two-way repeated-measures ANOVAs to test for the role of valence and information manipulation, as well as their interaction. The direction of changes was analyzed by follow-up t-tests. In particular, one-sample t-tests were used when comparing data to a reference value (e.g., guessing level: 50%), and paired t-tests were used to compared responses across different conditions (e.g., gain vs. loss) or different measures (e.g., averaged learning performance vs. averaged confidence).

The relationship between confidence and reaction time was analyzed by running a correlation analysis (Pearson's R). Using these correlation coefficients from all participants, we then tested if the coefficients were significantly larger than zero using one-sample t-test. All statistical analyses were performed using MatlabR2015a® (MathWorks) and its build-in functions (i.e., one-sample t-test: ttest; paired t-test: ttest2; repeated ANOVA: anovan; Pearson's correlation: corr).

4.2.5 Computational modelling

Participants' choices from both learning task and transfer task are fitted with reinforcement learning models (RL models). Similar to the approach we used in the Chapter2 (Ting et al., 2020a), we built up a nested model-space containing six increasingly complex RL models: RW, ABS, REL, REL_{α}, REL_{w}, $REL_{\alpha,w}$ (**Table 4.1A**). RW model referred to Rescorla-Wagner model, ABS and REL referred to ABSOLUTE and RELATIVE. The RELATIVE model was extended and assumed that other sources of information were integrated during learning. The first extension of REL was REL_{w}, which included weight parameter for counterfactual outcome in partial information condition. The second extension of REL was REL_{α}, which updated confirming and deconfriming information separately (Palminteri, Lefebvre, et al., 2017). The third extension of REL was the combination of REL_{w} and REL_{α} (**Table 4.1A**).

Models within model space updated chosen option value $Q(s,c)$ through delta-rule function at trials *t* under situation *s*. The unchosen option $Q(s,u)$ was updated with the same approach in REL models. Moreover, REL models updated chosen and unchosen option separately with α_c and α_u as follow:

$$Q_{t+1}(s,c) = Q_t(s,c) + \alpha_c \times \delta_c$$
$$Q_{t+1}(s,u) = Q_t(s,u) + \alpha_u \times \delta_u$$

Previous studies have shown that confirmation bias might involve in learning process (Palminteri, Lefebvre, et al., 2017). To capture this bias, REL_{α} and $REL_{\alpha,w}$ contained two learning rates: α^+ and α^-. In particular, α^+ served as confirmation updating, which updated positive prediction error for chosen option (i.e., confirming chosen option would lead to better outcome) and updated negative prediction error for unchosen options (i.e., confirming unchosen option would lead to worse outcome). By contrast, α^- was used to update deconfirming outcome (i.e., negative prediction error for chosen options and positive prediction error for unchosen options).

$$Chosen\ option \begin{cases} Q_{t+1}(s,c) = Q_t(s,c) + \alpha^+ \times \delta_{c,t}, & if\ \delta_{c,t} > 0 \\ Q_{t+1}(s,c) = Q_t(s,c) + \alpha^- \times \delta_{c,t}, & if\ \delta_{c,t} < 0 \end{cases}$$

$$Unchosen\ option \begin{cases} Q_{t+1}(s,u) = Q_t(s,u) + \alpha^+ \times \delta_{u,t}, & if\ \delta_{u,t} < 0 \\ Q_{t+1}(s,u) = Q_t(s,u) + \alpha^- \times \delta_{u,t}, & if\ \delta_{u,t} > 0 \end{cases}$$

Prediction error for chosen option δ_c was calculated as difference between chosen outcome R_c and estimated chosen option value $Q(s,c)$. On the other hand, prediction error for unchosen option δ_u is calculated as difference of unchosen outcome R_U and estimated unchosen option value.

The prediction error was corrected by considering context value V in any REL models. That is, factual outcome is subtracted by context value and then used to compute prediction error, such as:

$$\delta_c = R_c - V(s) - Q_t(s,c)$$
$$\delta_u = R_u - V(s) - Q_t(s,u)$$

The context value V was also updated through delta-rule with its own prediction error δ_V:

$$\mathrm{V}_{t+1}(\mathrm{s}) = \mathrm{V}_t(\mathrm{s}) + \alpha_V \times \delta_V$$

, where α_V was the context value learning rate and δ_V was a prediction error calculated as follow:

$\delta_V = (\mathrm{R}_{C,t} + \neg R_t(s))/2 - V_t(s)$, if partial information condition
$\delta_V = (\mathrm{R}_{C,t} + \mathrm{R}_{U,t})/2 - V_t(s)$, if complete information condition

When outcome for forgone option was not available in context s (i.e., partial information condition), we assumed participants might make inference $\neg R_t(s)$ for forgone option given chosen outcome and sign of context value:

$$\neg R_t(s) = \begin{cases} 1 \; if \; |R_t(s)| = 0.1 \; and \; V_t(s) > 0 \\ -1 \; if \; |R_t(s)| = 0.1 \; and \; V_t(s) < 0 \\ 0.1 \; if \; |R_t(s)| = 1 \; and \; V_t(s) > 0 \\ -0.1 \; if \; |R_t(s)| = 1 \; o \; and \; V_t(s) < 0 \end{cases}$$

Therefore $\neg R_t$ captured the fact that participants inferred that the non-selected cue was associated with the complementary outcome to the one they actually received. The formulation of $\neg R_t$ depended on the context value $V_t(s)$ because context values have to be disambiguated (i.e. gain or loss context) before participants can infer the complementary outcome. However, it was likely that $\neg R_t$ was not fully used in the learning updating. To account for this possibility, $\neg R_t$ was multiplied by weight parameter w when it was used to update context value:

$$\delta_V = (R_t(s) + \mathrm{w}(\neg R_t(s)))/2 - V_t(s)$$

Finally, the probability of choosing better option *A* in learning task was computed based on the softmax function:

$$\mathrm{P}_{learning}(s, A) = (1 + \exp\left(\beta\left(Q_t(s, A) - \left(Q_t(s, B)\right)\right)\right)^{-1}$$

The same softmax function and the same inversed temperature parameter β were applied for transfer task to compute the probability of choosing better option *A* (i.e., the one with higher expected value):

$$\mathrm{P}_{transfer}(s, A) = (1 + \exp\left(\beta\left(Q_{end}(s, A) - \left(Q_{end}(s, B)\right)\right)\right)^{-1}$$

, where Q_{end} referred to the option value in the end of learning trial.

The inversed temperature parameter β represented the randomness of decision given option values. That is, the larger β indicated the choice was made given expected value. In contrast, the lower β indicated the choice was deviated from expected value. Both option values *Qs* and context values *Vs* in each condition are set at 0 in the beginning, and then are updated given choice, feedback and condition.

4.2.6 Model optimization and comparison

The parameters θ in each model M were optimized via procedure of minimizing negative logarithm of the posterior probability (n*LPP*):

$$nLPP = -\log\big(P(\theta_M|D,M)\big) \propto -\log\big(P(D|M,\theta_M)\big) - \log\big(P(\theta_M|M)\big)$$

$P(D|M,\theta_M)$ refers to the likelihood of choice D (i.e., choosing option with higher expected value) given current model M and its parameters θ_M. On the other hand, $P(\theta_M|M)$ is the likelihood of getting θ_M within prior probability of the parameters. The prior distribution of learning rates and weight were defined as beta distributions with two parameters: a = b= 1.1. The prior distribution of inverse temperature parameter β was defined as gamma distribution with two parameters: a =1.2 and b =5 (Daw et al., 2011). The modelling modeling was conducted using Matlab 2015a. The estimated rating (Q) was initialized as the actual rating in the first trial in each block. All the parameters were initialized from random starting points selecting from the certain ranges (i.e., 0< α<1; 0< w<1; 0< β <Infinite) and were then optimized using Matlab's *fmincon*.

The lower LPP indicates the model can explain data better, however, n*LPP* doesn't take model's complexity into consideration. To address this issue, we calculated AIC, BIC and Laplace approximation to the model evidence (*LAME*), which penalized model's complexity (i.e., number of parameters). Three model comparison criteria for each model were computed as followed:

$$\mathrm{BIC} = -nLPP - \frac{df}{2}\log(n)$$

,

$$\mathrm{AIC} = -nLPP\mathrm{P} - df$$

and

$$\mathrm{LAME} = -nLPP + \frac{df}{2}\log(2\pi) - \frac{1}{2}\log|H|$$

, where n was number of trials, *df* was determined as number of parameters and $|H|$ was the determination of the Hessian. Again, these criteria were computed at individual level.

The winning model in the model comparison doesn't imply the model can generate a certain data set at group-level random-effect analysis. To assess the model's reliability, we then fed LAME (from each subject in each model) to mbb-vb-toolbox (http://mbb-team.github.io/VBA-toolbox/; Daunizeau et al., 2014). This toolbox performs Bayesian model selection procedure and estimates two indicators: the expected frequencies (*EF*) and the exceedance probability (XP) for each model. Specifically, the expected frequency *EF* of a model quantifies the probability that the

model generated the data for any randomly selected subject. Note that the EF should be higher than chance level given number of models in the model space. Exceedance probability (XP), on the other hand, quantified the belief that the model is more likely than all the other models of the model-space.

Different parameters might be able to create the same behavioral patterns and reduced model's identifiability (Palminteri, Wyart, et al., 2017). To ensure models within model space have good ability to replicate behavioral patterns, we simulated 500 synthetic data for each model and assessed both ability of parameter recovery and ability of model recovery. Specifically, we tested (1) if the parameters used to simulate data significantly correlate with estimated parameter, and (2) if simulated data can be only explained by the model that generated them. We applied correlation analysis to assess the relationship between simulated parameters and estimated parameters. Larger correlation coefficient (R) and explain variance (R^2) indicated better parameter recovery. The model identifiability was assessed by Bayesian Model comparison using the mbb-vb-toolbox.

4.2.7 fMRI

4.2.7.1 fMRI preprocessing

The functional images were preprocessed using SPM12 (Wellcome Department of Imaging Neuroscience, London) with the following steps: realignment and unwarp, slice timing correction, co-registration, segmenting anatomical images, normalization and smoothing. To correct for potential head movement during functional images collection, all functional volumes (from three runs) were realigned to the first volume in the first run and were unwarped with collected field maps. Next, slice timing correction was performed. To improve the quality of following normalization, the mean functional (the output from step1) and anatomical images were coregistered. Afterwards, the anatomical image from each subject was segmented into six images (i.e., grey matter, white matter, cerebrospinal fluid, fat tissue and air) using nonlinear deformation fields and SPM12's Tissue Probability Maps (TPMs). To ensure images from different individuals are comparable, all images were normalized to the Montreal Neurological Institute T1 template (i.e., MNI152) using forward deformation fields from the segmentation output. Finally, the EPI images were normalized and smoothed with a full width half maximum Gaussian kernel of 6-mm (2 times of voxel size of functional images) full-width at half maximum (FWHM) isotropic Gaussian kernel.

4.2.7.1 fMRI analysis: whole-brain analyses

One of the main purposes of the present chapter was to investigate the general and condition-based neural mechanisms of confidence and outcome evaluation. To address this goal, we specified a general linear model (GLM) to analyze imaging data at both whole-brain and region of interest (ROI) level. As we were mainly interested in mechanisms of valuation in each condition, we divided symbol onset and outcome onset into four conditions (i.e., GP, LP, GC, LC) and created a total of eight corresponding regressors (four for symbol phase, four for outcome phase). Two more event onsets, choice phase and rating phase, were included without division. These event-related regressors were modeled using stick (delta) functions. Five types of parametric modulators were included in the GLM: (1) confidence ratings for each condition-specific symbol onset, (2) reaction time for choice onset, (3) reaction time and distance between initial and final rating point for rating onset and, (4) receiving outcome (code as 1 and 0 for relatively good and relatively bad outcome, respectively) for each condition-specific outcome phase. Parametric modulators were z-scored to ensure results from different condition and regressors were comparable. To remove motion artifact and to improve the quality of fMRI results, the GLM also contained six realignment parameters, which were created during preprocessing. Consequently, a total of 27 regressors were included for each run.

In order to identify the brain regions that generally encode confidence ratings regardless of experimental manipulations, we performed a contrast grouping effects of confidence rating from all conditions. We were also interested in the effect of experimental manipulations (i.e., outcome valence and information) and their interaction on confidence encoding, which we tested by contrasting the parametric modulator for confidence across conditions using the following interaction contrasts based on our factorial design: (GP+GC)>(LP+LC), (GC+LC)>(GP+LP) and (GC-GP)>(LP-LC). The same contrasts were applied to the parametric modulator reflecting

good/bad outcomes per condition. The GLM and contrasts were performed at the individual-level (first-level). Subsequently, results from all participants were taken to the group-level random effect analysis (second-level) using simple t-test with both exploratory whole-brain cluster-defining height threshold at uncorrected $p<.001$ and family-wise error (FWE)-corrected threshold of $p<.05$.

4.2.7.3 fMRI analysis: ROI analyses

The confidence-encoding and outcome-encoding brain regions we identified from the whole brain analysis were used in follow-up ROI analyses. Specifically, we created masks (not spheres) based on the clusters we identified from the contrast of parametric modulation of confidence rating and the contrast of parametric modulation of outcome using marsbar (Brett et al., 2002). These masks served as regions of interest (ROI) and data from these regions was extracted using either spm build-in function: spm_get_data.m or rfxplot (Gläscher, 2009) for further condition-based analyses. Two participants were excluded, one has significant dropout in the functional images and one has invariable responses on confidence rating leading to difficulty on performing followed-up t-contrast, and total 38 participants were analyzed for the fMRI results.

The ROI was performed to investigate condition-dependent confidence and outcome encoding. Using the original GLM, the regression coefficients were extracted and averaged from each mask for each condition. Note that the underlying test that was used to create ROIs, a main effect of parametric modulation independent of experimental conditions, is orthogonal to the follow-up tests performed on data extracted from ROIs (condition-specific effects), therefore these analyses are not circular (Kriegeskorte et al., 2009). The condition-based data were analyzed by simple t-tests to assess if the parametric modulator for confidence reflects differential correlations with the BOLD signal across the experimental conditions. We also performed two-way repeated-measures ANOVA to test the main effect of experimental manipulations and their interaction.

4.3 Results

4.3.1 Learning performance and confidence

In total, 40 participants completed the learning task in the MRI scanner. Here we successfully replicate the results from last chapter. First of all, the average probability of choosing the better symbol was above guessing level ($t_{39} = 17.78$; $p < .001$; **Table C.1.1**), which indicated participants were able to learn the better symbols from learning task. Secondly, replicating the results from Fontanesi et al., 2019; Palminteri et al., 2015, a two-way repeated-measures ANOVA showed that learning performance was affected by the information ($F_{1,39} = 22.05$, $p < .001$), but not by the outcome valence ($F_{1,39} = 0.00$, $p = .9666$). No interaction between information and outcome valence was found ($F_{1,39} = 0.01$, $p = .9056$). With the same ANOVA analysis, we replicated findings from last chapter (Ting et al., 2020b) and (Lebreton et al., 2019) and found the significant main effect of outcome valence on both confidence ($F_{1,39}=36.56$, $p<.001$) and reaction time ($F_{1,39} = 4.77$, $p = .0350$). However, the main effect of information manipulation ($F_{1,39}= 6.76$, $p =.0131$) and their interaction ($F_{1,39}= 9.62$, $p =.0036$) were found in confidence ratings only (**Fig 4.2A**; **Table C.1.2**). The *post hoc* results showed that the differences in confidence between gain and loss was larger in the partial information condition ($t_{39} = 6.93$, $p = 2.68\times10^{-8}$) compared to complete information condition ($t_{39} = 4.55$, $p = 5.08\times10^{-5}$) (**Figure 4.2B**).

Thirdly, we replicated the negative correlation between confidence and reaction time found by Chapter 3 (Ting et al., 2020b) at both group-level and condition-level (**Table C.1.3**). Nevertheless, a robust regression showed that the valence-induced confidence changes and the valence-induced RT changes were not correlated as the slope was not significant ($\beta = -0.01\pm 0.01$, $p = .339$). Moreover, the significant and positive intercept ($\beta = 5.02 \pm 0.84$; $p <.001$) from the robust regression implied that the confidence bias might be observed when RT bias was absent (**Table C.1.4**). These results are in line with our observation from chapter 3 that the valence-induced changes on confidence and on RTs can be partially dissociable.

4.3.2 The condition-based confidence bias

As outcome valence significantly affected confidence ratings but not learning performance, it is likely that confidence deviated from actual performance and resulted in a *confidence bias* (i.e., a discrepancy between confidence and actual performance). We directly computed such confidence bias by contrasting individuals' average confidence ratings with their actual average probabilities of choosing the better symbol. A positive value represents overconfidence and a negative value represents underconfidence. We found that the difference between confidence and performance on average was non-significant ($t_{1,39} = 0.1883$, $p = .8516$). To further understand whether confidence bias was modulated by valence and information manipulation, we performed two-way repeated ANOVA on confidence bias. Replicating previous finding (Lebreton et al., 2019), we found significant main effect of outcome valence ($F_{1,39} = 12.28$, $p < .001$) and information ($F_{1,39} = 14.42$, $p <.001$) on confidence bias, but not interaction ($F_{1,39} = 0.58$, $p =.4506$). The *post hoc* analysis showed that the confidence was decreased when more information was available and resulted in mitigated overconfidence in gain conditions (Gain complete> Gain partial: $t_{49} = -2.70$, $p =.0100$)

and enhanced underconfidence in loss conditions (Loss complete> Loss partial:t_{49} = -2.90, p =.006).

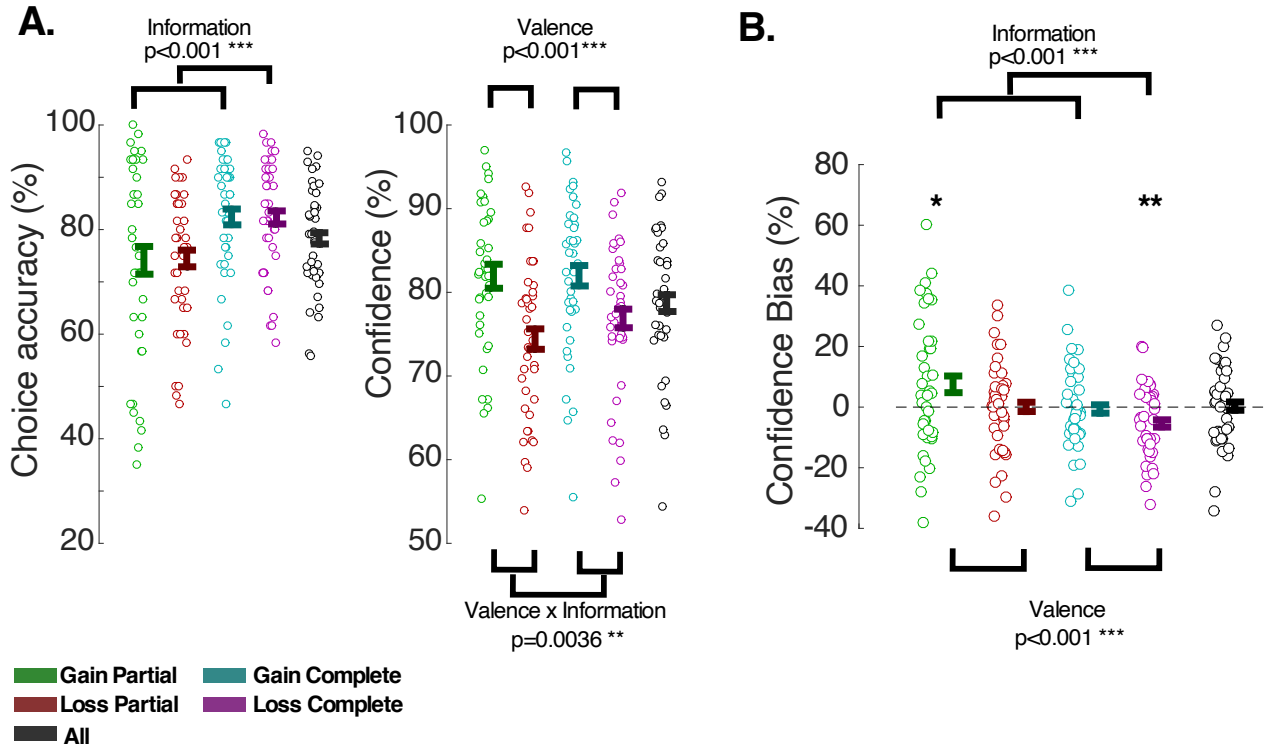


Figure 4.2. The effect of valence and information on learning performance, confidence and confidence bias. Using **(A) Performance.** Average choice accuracy at both individua-level and group-level across conditions. Choice accuracy was determined as choosing symbol associated with higher expected value in each condition. **(B) Confidence.** Average confidence level at individua-level (dots) and group-level (error bars) across conditions. **(C) Confidence bias.** Average confidence bias at individua-level (dots) and group-level (error bars) across conditions. Confidence bias was measured as the differences between averaged confidence and averaged probability of correctness. The start signs represented the results from simple t-test, which tested whether confidence bias was significantly different from zero. Empty dots with colored edges represent individual data points. The colored error bars displayed on the side of scatter plots represent the mean±SEM. The p-value above and below the figure showed the results of two-way repeated-measured ANOVA.
~p < .1; *p < .05; **p < .01; ***p < .001

4.3.3 Context value and confirmation bias in reinforcement learning model

The valence-induced changes on confidence suggested that context value and outcome valence were integrated to option value during learning. Moreover, we found that participants tend to make inferences for unchosen outcomes when partial information (i.e., chosen outcome) was given in Chapter 2. Replicating findings from Chapter 2 (Ting et al., 2020a), the model comparison across five model selection criteria showed that the wining model was $REL_{\alpha,w}$ (**Table 4.1B**). In addition to relative model comparison criteria, we also applied parameter and model recovery tests to verify the models in model space were reliable (Palminteri, Wyart, et al., 2017) (**Figure C.1.1**). The results indicated that both confirmation learning rates (α^+) and weight for forgone outcome (w) in partial information condition were involved in the learning process. Moreover, the positive learning rate was significantly higher than negative learning rate ($t_{39}= 7.66$, $p < 2.72\times10^{-9}$; **Table C.1.5A**), which replicated the learning asymmetry we found in Chapter 2 and in previous studies (Lefebvre et al., 2017; Palminteri, Lefebvre, et al., 2017). The results indicated that the model we developed is robust when the task was conducted with and without emotion manipulation.

The parameters in the winning model reflect the learning strategy (or bias). For example, inverse temperature parameter was usually associated with exploitation-exploration tradeoff and confirmation learning rate was associated with confirmation bias. In order to understand if confidence bias (i.e., the discrepancy between actual learning performance and average confidence rating) was linked to learning strategy (or bias), we conducted an exploratory analysis to investigate the relationship between confidence bias and learning parameters (**Table C.1.5B**). The correlation analyses showed that the average confidence bias negatively correlated with the inverse temperature parameter ($r = -0.58$, $p = .0001$), which implied that overconfidence was associated with more exploitation (i.e., relied on expected values to make choice) and underconfidence was associated with more exploration (i.e., less reliance on expected values). The result was consistent with a recent study showing the correlation between confidence and exploitation/exploration tradeoff (Boldt et al., 2017).

Table 4.1. Model space and results of model comparison.

A. Models in the model space and corresponding parameters

Model	Learning	Context	Decision
RW	α_c		β
ABS	α_c α_u		β
REL	α_c, α_u	α_v	β
REL_{α}	α^+, α^-	α_v	β
REL_w	α_c, α_u	α_v, w	β
$REL_{\alpha,w}$	α^+, α^-	α_v, w	β

Models in the model space and corresponding parameters. The learning-related parameters included learning rates updating option value; the context-related parameters included learning rates updating context value and weighting on making inference for unchosen outcome; decision-related parameter was inverse temperature parameter, which represented the randomness of decision given estimated option values.

B. Model Comparison

	DF	LAME	AIC/2	BIC/2	EF	XP
RW	2	196.8±30.1	231.7±14.1	235.2±14.1	0.00	0
ABS	3	189.6±33.0	187.6±30.3	192.8±30.3	0.02	0
REL	4	169.4±33.6	166.7±33.4	173.9±33.4	0.00	0
REL_{α}	4	161.3±33.0	156.7±32.7	163.7±32.7	0.37	0.21
REL_w	5	166.6±34.8	162.4±35.2	171.1±35.2	0.08	0
$REL_{\alpha,w}$	**5**	**158.6±33.3**	**154.3±33.0**	**163.0±33.0**	**0.50**	**0.78**

The results of model comparison. All model comparison criteria were calculated based on LPP. The winning model ($REL_{\alpha,w}$) was highlighted with bold font. EF and XP were calculated using mbb-vb-toolbox (Daunizeau et al., 2014).

AIC: Akaike information criteria; **BIC**: Bayesian information criteria; **LAME**: Laplace approximation to the model evidence; **EF**: Efficient frequency; **XP**: Exceedance probability.

4.3.4 Outcome-related signal was modulated by valence manipulation

The winning model in the current study suggests that the learning process consisted of two important elements during outcome presentation. Firstly, option value was updated by comparing actual outcome and expected outcome. Secondly, the context value was learnt and used to calibrate outcome. To ensure these elements were processed at the neural level, we analyzed brain activity from fMRI data and focused on the outcome phase (**Appendix C.2: Table C.2.1**). We found that outcome-related signal (i.e., parametric modulator of good and bad outcomes in each condition) was generally and positively encoded in the neural network including bilateral ventral striatum (VS), putamen, caudate and hippocampus (k =46, uncorrected threshold of p<.001; **Table C.2.2**). Using a conservative threshold, only right ventral striatum (rVS) survived (k =46, FWE-corrected p<.05).

To further test whether the outcome-related signal in rVS was condition-dependent, we created a ROI mask for rVS and separately extracted regressor coefficients from the modulator of relatively good and bad outcomes for each condition. We found that these coefficients were significantly higher than 0, indicating relatively good outcome increased activity in rVS across four conditions (ps < .01; **Figure 4.3**). We then fed these coefficients into a two-way repeated-measures ANOVA to test if our manipulation significantly modulated the way of outcome processing. We only found a partially significant main effect of valence ($F_{1,37}$= 3.8104, p = .059), which implied that the good outcomes in the gain context may induced higher activity in rVS.

We found another neural network including right Insula, dorsal medial prefrontal cortex (dmPFC) and bilateral Inferior frontal gyrus (IFG), which negatively encoded relatively good vs. bad outcomes (k= 46, uncorrected threshold of p<.001; no survival under p<.05, FWE-corrected; **Table C.2.2**). Again, we tested if these signals were condition-dependent via followed-up ROI analyses given the identified clusters. We found that modulator of relatively good and bad outcomes negatively correlated with brain activity in Insula and dmPFC across four conditions, indicating that these regions preferentially track relatively bad outcomes. The two-way repeated ANOVA results showed that a main effect of outcome valence in Insula ($F_{1,37}$ = 5.57; p = .0237) and bilateral IFG ($F_{1,37}$ = 5.57; $p = 2.87\times10^{-4}$). The main effect of valence on outcome-related signal in Insula was driven by lower coefficients in loss contexts (t_{37}= 2.36; p = .0237). In contrast, the valence effect on outcome-related signal in bilateral IFG was driven by lower coefficients in gain contexts (t_{37}= -4.01; $p = 2.87\times10^{-4}$). With the same ROI analysis, we didn't find main effect nor interaction in dmPFC. Together, these results suggested that outcome evaluation was modulated by context as relatively bad outcome in loss contexts decreased Insula activity more than in gain context (Palminteri et al., 2015).

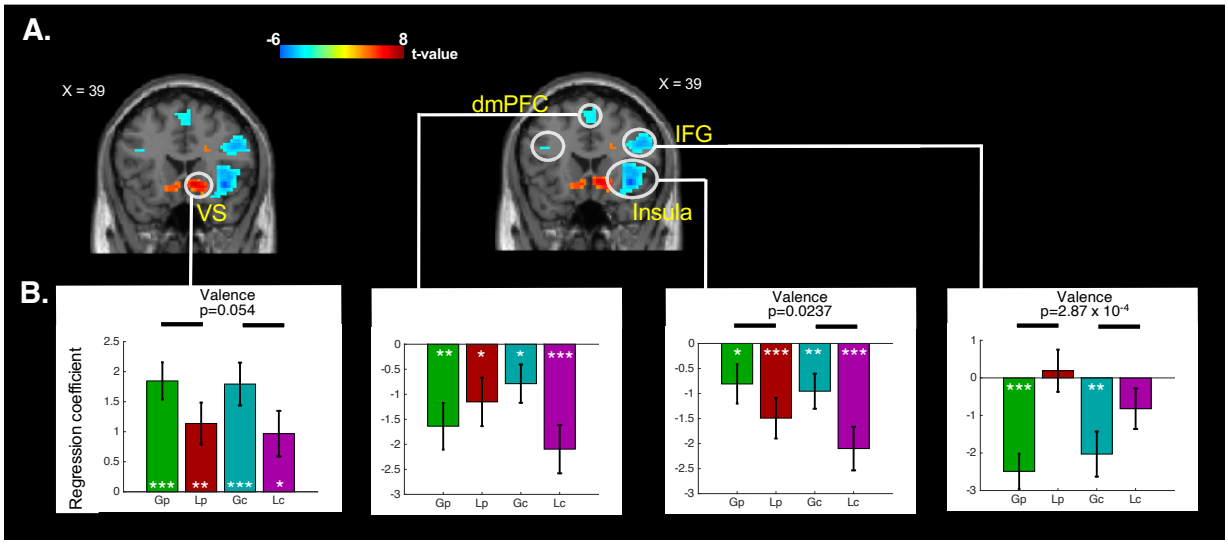


Figure 4.3. The neural evidence of valence effect on outcome evaluation. (A) Outcome (good/bad) was positively correlated with activation in many regions (p<.001, uncorrected; **Table C.2.2**), but only VS (peak at xyz = [9 14 -10]) was survived under cluster-defining height threshold of FWE-corrected p<.05. The negative correlation between confidence ratings and activation in Insula (peak at xyz =[39 20 -10]), dmPFC (peak at xyz = [6 35 47]) and bilateral IFG (peak at xyz = [48 17 29] and at xyz = [-45 5 29]). The significant clusters under uncorrected threshold of p<.001 were illustrated on xjview's single T1 template with colors between red (positive correlation) to blue (negative correlation). **(B)** Bars represented the regression coefficients (y-axis) extracted from the clusters of VS, insula dmPFC and bilateral IFG for each condition. Different colors were used to represents dataset from different conditions (x-axis). The error bar on each bar represented mean ±SEM. We also tested if regression coefficient was significantly different from zero using simple t-test. The white star signs were used to illustrate the results of t-test.
Gp: gain/partial; **Lp**: loss/partial; **Gc**: gain/complete; **Lc**: loss/complete
~p < .1; *p < .05; **p < .01; ***p < .001

4.3.5 The neural mechanisms of confidence formation at symbol presentation

The previous analyses indicated that the neural network devoted to outcome evaluation was modulated by the outcome valence manipulation. Previous studies suggest that confidence rating was automatically integrated to the value-based decision, we therefore performed a whole-brain analysis using the parametric modulator of general confidence (i.e., [Gp+Lp+Gc+Lc]) during symbol presentation to investigate how confidence is neurally integrated during the decision moment. We found a condition-general neural network in brain value system that positively encoded confidence in vmPFC and negatively encoded confidence in caudate and Bilateral Insula (k=41, uncorrected threshold of p<.001; **Table C.2.2**). With the same threshold, no other contrasts (i.e., interaction and Gain>Loss) yielded effects. These results indicate that confidence was generally encoded in the same brain network regardless of affective contexts.

4.3.6 Context-based confidence encoding

To test how each condition contributed to the identified condition-general neural network, we extracted regressor coefficients for the parametric modulator of confidence from each condition.

This condition-based ROI analysis showed that reported confidence ratings were not equally processed in gain and loss contexts, even though the direction of encoding was the same. Specifically, the positive correlation between confidence and activity in vmPFC was not significant in loss/complete condition (p= .0684) the negative correlation between confidence and activity in caudate and bilateral insula were not significant in loss/partial condition (p= .2683; **Figure 4.4B**).

To ensure confidence was selectively encoded or not encoded for either loss/partial or loss/complete conditions, we conducted two more contrasts: [(Gp+Lp+Gc)>Lc] and [(Gp+Gc+Lc)>Lp]. We replicated the results above showing significant correlation in vmPFC and right Insula (cluster-identifying uncorrected threshold of p<.001), which mainly driven by insignificant correlation in loss/complete and loss/partial condition, respectively. We didn't find other brain regions under two additional contrasts, confirming that no more brain regions were involved in confidence processing in loss contexts during symbol presentation.

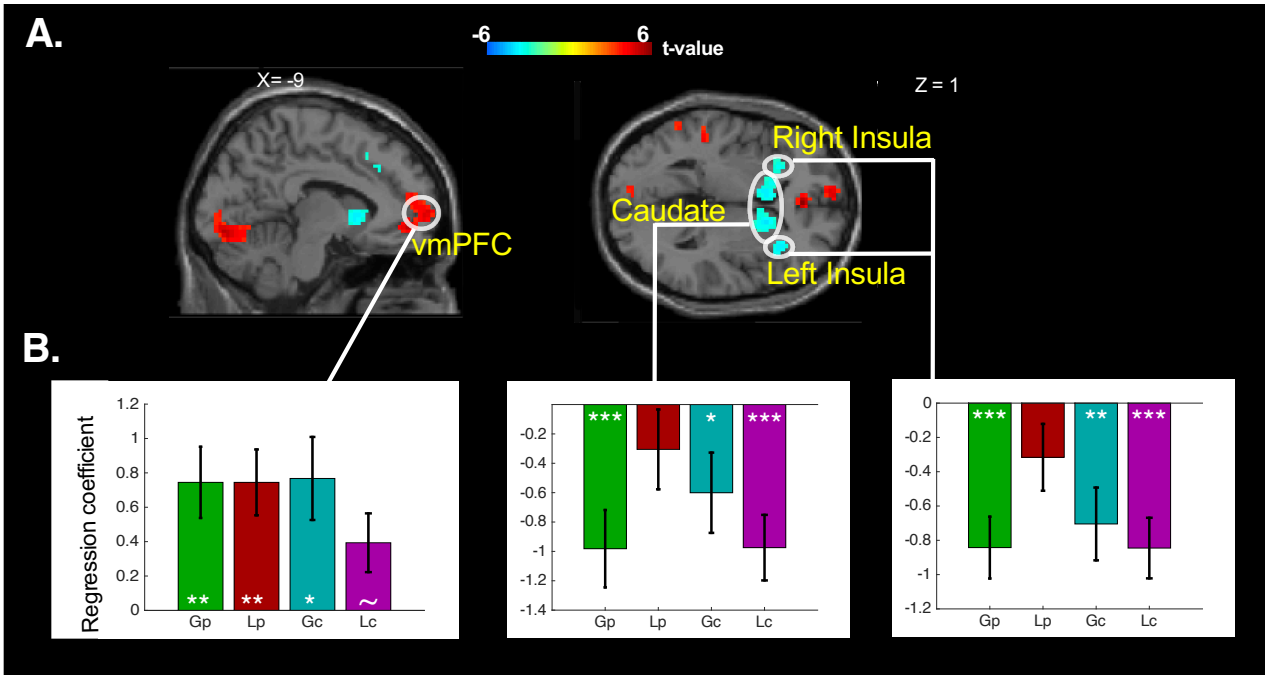


Figure 4.4. The neural confidence signals. (A) Confidence rating positively correlated with activation in vmPFC (peak at xyz = [0 41 2]), and negatively correlated with activation in bilateral insula (peak at xyz = [33 23 -4] and at xyz = [-30 23 -4]) and caudate (peak at xyz = [-9 11 2]). The significant clusters under uncorrected threshold of p<.001 were illustrated on xjview's single T1 template with colors between red (positive correlation) to blue (negative correlation). **(B)** Bars represented the regression coefficients (y-axis) extracted from the clusters of vmPFC, bilateral insula and caudate for each condition. Different colors were used to represents dataset from different conditions (x-axis). The error bar on each bar represented mean ±SEM. We also tested if regression coefficient was significantly different from zero using simple t-test. The white star signs were used to illustrate the results of t-tests.
Gp: gain/partial; **Lp**: loss/partial; **Gc**: gain/complete; **Lc**: loss/complete
~p < .1; *p < .05; **p < .01; ***p < .001

4.4 Discussion

The valence-induced changes on confidence in reinforcement learning have been well-demonstrated, however, the corresponding cognitive and neurological mechanisms underlying confidence formation and confidence bias under different affective contexts remain unclear. The present study is the first research investigating the neural mechanisms of confidence formation when outcome valence was clearly dissociated and when motor responses were orthogonalized with the option evaluation. We found that neural confidence signals were generally processed in a network consisting of ventral medial prefrontal cortex (vmPFC), bilateral insula and caudate. While the condition-based ROI analyses showed that the different conditions shared the same sign of correlation coefficients in each region of interest (**Figure 4.4**), the insignificant correlation in loss conditions implying that confidence ratings made in the negative environment were not fully formed or encoded in value-based neural network during symbol presentation (i.e., option evaluation).

Our results confirm and replicated that the affective context, which we manipulated via outcome valence, robustly influenced confidence at the behavioral level (Fontanesi et al., 2019; Lebreton et al., 2015, 2018, 2019; Palminteri et al., 2015; Ting et al., 2020b). Specifically, using a relatively large sample size (N=40) for fMRI studies (and compared to the experiments in the Chapter 3; Ting et al., 2020b), we observed significant differences in confidence ratings but equal learning performance between gain and loss contexts. Another robust finding was that valence-induced changes on confidence and reaction times (RTs) were dissociable, even when RTs were isolated from decision time via experimental design (see Methods). Outcome valence also altered the way of neural process of outcome. Replicating previous findings (Palminteri, Justo, Jauffret, Pavlicek, Dauta, Delmaire, Czernecki, Karachi, Capelle, Durr, et al., 2012; Palminteri & Pessiglione, 2017), the relatively good compared to bad outcomes, regardless of their valence, were encoded in well-documented value-based neural networks, including ventral striatum (VS), Insula, Inferior frontal gyrus (IFG) and dorsal medial prefrontal cortex (dmPFC). The ROI-analyses further demonstrated a significant effect of outcome valence on level of encoding, which suggested that the affective contexts enhanced signals of relative good and bad outcomes. Contrary to Palminteri et al., 2015, we failed to find a main effect of information nor an interaction. One possible reason are task differences, such that the worse outcome in the gain condition, and the better outcome in the loss condition still carried a value of 10 cents instead of 0 cents. Therefore, participants were less likely to be confused by 0 and could evaluate the context's "valence" even when they received the outcome with the lower numeric value. This possible reason was evident by the modeling result as the winning model suggested participants tended to make inference for the foregone outcome in partial information conditions.

The reported confidence was generally processed in a network consisting of vmPFC, bilateral insula and caudate. Specifically, during the presentation of the choice options, i.e. before an explicit requirement for communicating the choice and before stating one's confidence, confidence was generally processed in the same networks regardless of outcome valence and information manipulations. Confidence was positively encoded in vmPFC (**Figure 4.4**; **Table C.2.2**), which replicated previous findings (Bang & Fleming, 2018; Lebreton et al., 2015; Wittmann et al., 2016). Moreover, we also found that memory-related brain region: medial temporal lobe (MTL) encoded

confidence. The result not only replicated findings from Chua et al. (2014) and Wittmann et al. (2016) but also suggested that history of choices and outcome were involved in the learning task. On the other hand, we found that confidence was negatively encoded in another neural network including bilateral insula, caudate and dmPFC (**Figure 4.4**; **Table C.2.2**), which have been associated with risk evaluation and negative events (Palminteri, Justo, Jauffret, Pavlicek, Dauta, Delmaire, Czernecki, Karachi, Capelle, Durr, et al., 2012; Palminteri & Pessiglione, 2017). Our results are consistent with a recent study showing negative correlation between confidence and activity in these regions in perceptual decision making (Hebart et al., 2016). Altogether, the results indicate that reported confidence was formed and encoded before the actual choice was communicated and before confidence was explicitly rated (Lebreton et al., 2015). This interpretation is not only in line with our conclusion in the last chapter and recent studies arguing confidence is not mere consequence of performance or reaction time, but also studies arguing that confidence plays an important role on speed-accuracy tradeoff as longer time might be required when individual is not confident (Desender, Van Opstal, et al., 2017; Dotan et al., 2018a).

The follow-up ROI-analyses demonstrated how confidence ratings were processed in each specific condition. We found confidence ratings in gain contexts were consistently (positively/negatively) encoded regardless of information manipulation. By contrast, the correlation between confidence ratings and neural activity was modulated by information in loss contexts. On the one hand, confidence ratings from loss/complete condition only showed a trend of confidence encoding in vmPFC, the regions that positively processed confidence on average. On the other hand, confidence ratings from the loss/partial condition were not significantly encoded in the Insula and caudate, the regions that negatively processed confidence on average. The results implied that reported confidence was formed and encoded in gain contexts but not in loss contexts during symbol presentation. This idea was supported by the results showing no other brain regions specifically encode confidence ratings made in any loss contexts (see *Context-based confidence encoding*).

Although the results suggested the confidence were generally formed and encoded during symbol presentation (i.e., option evaluation), we cannot rule out the possibility that confidence is continuously accumulating evidences after option evaluation (post-decisional period; Fleming & Daw, 2017; Pleskac & Busemeyer, 2007) because of two reasons. Firstly, subregions of prefrontal cortex play different roles on confidence formation at different decision time points. For example, vmPFC was associated with uncertainty about estimated option during the option presentation period (i.e., when the confidence rating is not required) and rlPFC was associated with metacognitive report, especially when choice is made and when reporting confidence (De Martino et al., 2013; Fleming et al., 2012; Lebreton et al., 2015). Therefore, the rlPFC is likely to integrate information and form the reported confidence ratings. As we focused on confidence processing during symbol presentation rather than confidence rating period, we only observed vmPFC encoded confidence. Secondly, mPFC integrated not only value-related information but also motor signals from basal ganglia (Haber and Knutson, 2009) or/and from motor systems (Cerliani et al., 2012). The correlation between confidence bias and exploration/exploitation tradeoff also implied that cognitive control might interact with confidence judgments. To further understand how confidence judgment was influenced by other sources of information, the further analyses are required to investigate how confidence was dynamically formed right after option evaluation using functional connectivity analyses. We believe the results will shed a light on explaining valence-

induced confidence changes and confidence bias, which is one of ultimate goals in the present fmri study.

There are two limitations in present study. Firstly, we were not able to explain why confidence deviated from actual performance, despites that investigating neural basis of over- and under-confidence is one of the main goals of present study. Nevertheless, our behavioral and fMRI results provided a possible mechanism to explain confidence bias. The behavioral results revealed that averaged confidence rating was significantly higher than averaged accuracy rate in gain/partial condition. The over-confidence bias was mitigated when more information was available (gain/complete condition) or when participants were learning to avoid loss (loss/partial condition). Little to no overconfidence in loss contexts was in line with affect-as-information theory, which suggested that negative events would facilitate deliberative thinking and reduce bias (Clore & Huntsinger, 2007). This result implied that negative contexts and more information might require longer time or different neural system to "correct" confidence. However, the underconfidence we observed in loss/complete condition implied that more negative information is likely to overcorrect the confidence. This phenomenon might be explained by our fMRI result showing an absence of significant correlations between vmPFC signal and confidence in the loss/complete information condition only.

The second limitation is that we were not able to track confidence accuracy trial-by-trial in the fMRI analyses. Given the notion of two-bandit instrumental learning task, the continuous variable of confidence rating was not comparable to binary choice. Consequently, we have limited access to identify the regions correcting confidence, like anterior prefrontal cortex (aPFC). The aPFC activity has been associated with metacognitive ability: higher activation in aPFC was linked to lower confidence bias (Molenberghs et al., 2016). Therefore, investigating how and what information is integrated and formed the final reported confidence, especially in the loss contexts, will provide valuable information to directly/indirectly explain condition-based confidence bias. To address this issue, the further functional connectivity analyses are required to uncover the dynamic process of confidence formation.

Summary

The impact of contexts on decision-making is observable in field and laboratory experiments. The mechanisms of context-induced changes on choice patterns have been widely investigated at both behavioral and neural levels. For instance, reference points in prospect theory explain why the same set of options is evaluated differently. However, it remains less clear how affect and prior performance are involved in decision formation over time. Reinforcement learning provides a framework to address this issue by quantifying this updating process. The present thesis combines a modified reinforcement learning task with methods from affective neuroscience, computational modelling and brain imaging to systematically investigate mechanisms underlying the effects of two affective framing manipulations (i.e., choice-irrelevant affect: incidental anxiety; choice-relevant affect: outcome valence) on belief (i.e., option value and confidence judgments) updating. Our results demonstrate that incidental anxiety and outcome valence have limited impact on learning *per se*. Importantly, the impact of outcome valence on reaction time and confidence judgments is robust. These findings not only imply that confidence, option value and context are updated simultaneously, but the robust valence-induced changes of confidence indicate metacognitive function might be more sensitive to affective framing manipulations. Besides the empirical evidence concerning the role of anxiety and outcome valence in reinforcement learning, we also established useful methodological and analytical approaches for future studies investigating emotion-cognition-metacognition interactions in the field of judgment and decision-making.
The main findings for each chapter are separately summarized as follows:

Chapter 2 investigates how and whether incidental anxiety influences instrumental learning, while addressing and improving upon several issues with prior research identified in a focused literature review. We used a rich within-subject design, featuring both a learning and a transfer phase, and two affective framing manipulations: environment (anxiety vs safe) and outcome valence (gains vs losses). In two variants (N = 2x50) of this experimental paradigm, incidental anxiety was induced by delivering unpredictable, aversive and performance-independent electric shock during learning task. Moreover, the anxiety induction was assessed by both questionnaire and physiological responses (i.e., skin conductance responses). A comprehensive modelling effort revealed that, irrespective of the effects of anxiety, individuals give more weight to positive than negative outcomes and tend to experience the omission of loss as a gain (and vice versa). However, in line with results from our targeted literature survey, isolating the specific computational effects of anxiety on learning *per se* proved to be challenging. Overall, our results suggest that learning mechanisms are more complex than traditionally presumed and raise important concerns about the robustness of the effects of anxiety previously identified in simple reinforcement-learning studies.

Chapter 3 focuses on the impact of choice-related affect (i.e., outcome valence: gain vs loss) on three common measurements: accuracy, confidence and reaction time (RT). While these three measurements are supposed to be highly correlated, it remains controversial whether valence-induced confidence and RT changes are dissociable. In order to address this issue, the goal of chapter 3 is to assess the presence of the valence-induced confidence bias in the absence of the RT bias. We conducted six variants of a learning task, attempted to disrupt the valence-induced motor bias effects by manipulating the mapping between decisions and actions and imposing constraints on response times (RTs). We observed both motor and confidence biases despite our disruption attempts, establishing that the effects of valence on motor and metacognitive responses

are very robust and replicable. Nonetheless, within- and between-individual inferences reveal that the confidence bias resists the disruption of the RT bias. Therefore, although concomitant in most cases, valence-induced motor and confidence biases seem to be partly dissociable. These results highlight new important mechanistic constraints that should be incorporated in learning models to jointly explain choice, reaction times and confidence.

Chapter 4 tests whether confidence formation in the brain is context-dependent or context-independent. We combined fMRI and a reinforcement learning task optimized in chapter 3, which dissociated gain and loss contexts and isolated motor responses from option evaluation. This combination enabled us to measure task-related brain activity and further identify the brain regions involved in confidence processing. In this fMRI study (N=40), we successfully replicated the impact of valence on confidence and dissociated valence-induced confidence increases and valence-induced RT slowing. Importantly, we found the brain value system was generally involved in encoding confidence during symbol presentation and no other brain regions specifically encoded confidence ratings made in gain or loss contexts. Nonetheless, the correlation between reported confidence and activation in these regions was consistently found in gain contexts only but not in loss contexts. These results indicate that (1) confidence was formed and encoded in general neural network before explicit decision-making, and (2) the reported confidence in loss contexts might not be fully processed during option evaluation.

Appendices

APPENDIX A: Supplementary Material for Chapter 2

A.1 Small-scale Literature Review

A.1.1 Research Questions

The current meta-analysis has two goals; both goals mainly focus on the impact of anxiety on learning. First question is, how anxiety influences learning, including its process, performance and post-learning performance? Second question is, whether anxiety effect on learning is valence-dependent?

A1.2 Inclusion/Exclusion Criteria

We systematically review literatures investigating effect of trait anxiety and induced stress/anxiety in healthy adult on reinforcement learning. In the current mini-review, learning is determined as a simple probabilistic action-outcome association, which is then comparable for our research questions and experimental design. That is, individual should actively make choice given two or more options and then receive outcome.

Accordingly, we excluded studies with simple go-nogo learning task, two-stage reinforcement learning (which is used to differentiate model-free and model-based learning, such as (Gillan et al., 2020; Otto et al., 2013) or simple stimuli-response conditioning (i.e. Fear conditioning). In order to control the confounding effect of risk attitude, the task with gambling component (e.g., BART: Balloon Analogue Risk Task, IGT: IOWA gambling task) are excluded. With these criteria, we then included studies that were published in peer review journals between 2000 and 2019.

The performance is measured as percentage of choosing better symbol or number of blocks to reach training criteria. One study has two measurement: percentage of choosing better symbol and percentage of correctly naming action-outcome association (i.e., Schwabe & Wolf, 2009), so we separately displayed the effect for each measurement as measurement1 and measurement2, respectively.

To answer our first question, the stress/anxiety effect on learning is determined as the differences in choice between safe and stressful/anxious conditions. To further address the second question, we also summarized the interaction of anxiety and valence or difficulty in performance.

Learning

Author(s) and year	Threat Induction	Outcome valence	Task	Sample size	Anxiety Main effect	Interactions with Anxiety
Berghorst et al., 2013 Group	ToS (Between-subject)	Gain & Loss	PSST	N = 95 (Female only)	=	= Difficulty
Berghorst et al., 2013 SubGroup	ToS (Between-subject)	Gain & Loss	PSST	N = 45 (Female only)	=	= Difficulty
Petzold et al., 2010	TSST (Within-subject)	Gain & Loss	PSST	N=23 (F/M= 11/ 12)	=	N/A
Lighthall et al., 2013	CPT (Between-subject)	Gain & Loss	PSST	N= 96 (F/M = 48/ 48)	=	= Difficulty
Cavanagh et al., 2011‡	TSST (Within-subject)	Gain & Loss	PSST	N = 50 (F/M= 26/ 24)	↑	N/A
Schwabe and Wolf 2009 Measurement 1	SECPT (Between-subject)	Gain or Neutral	Action-Outcome association	N = 60	=	N/A
Schwabe and Wolf 2009 Measurement 2 #	SECPT (Between-subject)	Gain or Neutral	Action-Outcome association	N = 60	↓ ***	N/A
Glienke, et al., 2015	CPT (Between-subject)	Gain & Loss	Action-Outcome association	N= 40 (Male only)	=	= Difficulty
Treadway et al., 2017†	MIST (Within)	Gain or Loss	Probabilistic RL	N= 40 (Female only)	↑ *	= Outcome valence
DeVido et al., 2009	Patient study (Between-subject)	Gain or Loss	DRPLT	N = 55 (GSP/ GAD/ HC = 20/16/19)	↓*	= Outcome valence = Difficulty
Stevens et al., 2014	Social anxiety (Between-subject)	Loss	Probabilistic RL	N = 40 (HSA/NHSA = 20/ 20)	↑**	= Angry face **↑neutral face**
Safra, Chevallier, & Palminteri, 2018	Trait anxiety (within-subject)	Gain & Loss	Probabilistic RL	N = 100 (F/M = 70/ 30)	N/A	= Social information
Abraham and Hermann, 2015	Trait anxiety (between-subject)	Gain & Loss	PSST (face outcome)	N = 62 (HSA/NHSA = 30/ 32)	N/A	↑~ Difficulty
Voegler et al., 2019	Patient study (Between-subject)	Gain & Loss	PSST	N = 64 (SAD/HC = 34/ 30)	=	= Difficulty

Post-Learning

Author(s) and year	Threat Induction	Outcome valence	Task	Sample size	Anxiety effect	Interactions with Anxiety
Berghorst et al., 2013 Group	ToS (Between-subject)	Gain & Loss	PSST	N = 95 (Female only)	=	= Outcome valence
Berghorst et al., 2013 SubGroup	ToS (Between-subject)	Gain & Loss	PSST	N = 45 (Female only)	=	**↓ Gain** **↑ Loss**
Petzold et al., 2010	TSST (Within-subject)	Gain & Loss	PSST	N=23 (F/M= 11/ 12)	N/A	= Gain **↓ Loss ***
Lighthall et al., 2013	CPT (Between-subject)	Gain & Loss	PSST	N= 96 (F/M = 48/ 48)	=	**↑ Gain *** = Loss
Cavanagh et al., 2011†‡	TSST (Within-subject)	Gain & Loss	PSST	N = 50 (F/M= 26/ 24)	↑	= Outcome valence
Cavanagh, et al., 2019 ℘	Trait anxiety (within-subject)	Gain & Loss	PSST	N=121 (HC/DEP = 75/ 46)	N/A	= Gain **↑* Loss**
Abraham and Hermann, 2015	Trait anxiety (between-subject)	Gain & Loss	PSST (face outcome)	N = 62 (HSA/NHSA = 30/ 32)	N/A	= Gain **↑ * Loss**
Voegler et al., 2019	Patient study (Between-subject)	Gain & Loss	PSST	N = 64 (SAD/HC = 34/ 30)	N/A	= Gain (with/without pressure) **↑ ** Loss (without pressure)** = Loss (with pressure)

Table A.1. Review effect of aversive events on learning and post-learning/transfer task. This table separately summarizes previous findings about the effect of anxiety on learning (upper panel) and post-learning/transfer task (lower panel). Both training session (learning) and test session (Post-learning) are included. In the training session, subjects are asked to learn to select the better symbol in each fixed pairs of symbols. Each symbol is associated with certain probability of gains and reciprocal probability of losses. For example, a symbol is associated with 70% of wining 1€ and 30 % of losing 1€. Following the same example, another symbol in the fixed pair will then be associated with 70% of losing 1€ and 30 % of wining 1€.
PSST: Probabilistic stimulus selection task. **ToS**: Threat of shock. **TSST**: Trier Social Stress Test. **CPT**: cold-pressor test. **SECPT**: Socially Evaluated Cold pressure test. **MIST**: Montreal Imaging Stress Task
GSP: Generalized Social Phobia. **GAD**: Generalized Anxiety Disorder. HC: Health control. TR: treatment condition. HSA: High socially anxious. NHSA: non-socially anxious controls.
SAD: social anxiety disorder
F: Female. **M**: Male. RL: Reinforcement learning

Yellow area indicates that trait/clinical measures of anxiety were employed instead of anxiety induction methods.
Bold font indicates significant main or interaction effects of anxiety. =: no detectable impact. ↑: improvement; ↓: deficit
~: 0.05<p<0.1; *: 0.01< p<0.05; **: 0.001< p< 0.01; ***: p< 0.001

measures explicit knowledge about action-outcome association, which deviates from other measurements.
† effects of stress in this study are likely confounded by practice effects.
‡ The effect was calculated by given table
℘Correlation result

A.2 Demographics and behavior results

Table A.2.1. Demographics and Questionnaire

		Combined	Experiment 1	Experiment 2
Gender (M/F)		51/49	26/24	25/25
Age		23.27±3.08	24.14±3.09	22.40±2.85
BAI		10.17±8.61	9.90±8.43	10.44±8.86
SCR (Threat vs. Safe)		0.054±0.06	0.09 ±0.06	0.01±0.04
Self-reported emotional state toward threat condition	**Anxiety**	3.53±1.87	4.33±2.01	2.82±1.40
	Fear	3.29±2.04	3.95±2.32	2.70±1.55
	Sadness	1.96±1.29	2.33±1.50	1.64±0.96
	Happiness	2.61±1.65	2.62±1.83	2.60±1.49
	Angry	2.31±1.77	3.02±2.06	1.68±1.15
	Negative feeling (Safe)	--	--	3.48±1.63
	Negative feeling (Threat)	--	--	4.68±1.42
	Calm feeling (Safe)	--	--	6.50±2.15
	Calm feeling (Threat)	--	--	4.86±1.78

Reported values correspond to mean±STD.
STD: standard deviation

Table A.2.2. Learning performance and preferences.

A. Learning task (correct rate (%))

	Combined (Mean±SEM).	**Experiment 1** (Mean±SEM).	**Experiment 2** (Mean±SEM).
Safe Gain	70.94±2.15	74.46±3.00	67.43±3.03
Safe Loss	72.75±1.12	71.61±1.53	73.89±1.64
Threat Gain	74.85±1.74	77.42±2.53	72.27±2.35
Threat Loss	72.89±1.12	72.99±1.73	72.80±1.45

Accuracy rates in each condition during the learning task.
SEM: standard error of the mean

B. Transfer task (Preference (%))

Anxiety	Valence	**Combined** (Mean±SEM).	**Experiment 1** (Mean±SEM).	**Experiment 2** (Mean±SEM).
Safe	Gain 75%	76.93±2.79	83.83±4.09	70.04±3.56
	Gain 25%	47.70±2.83	41.33±4.73	54.07±2.88
	Loss 25%	50.68±1.75	50.50±3.04	50.86±1.74
	Loss 75%	24.68±1.93	24.33±3.48	25.04±1.72
Threat	Gain 75%	83.39±2.04	87.67±3.37	79.11±2.18
	Gain 25%	42.58±2.56	39.67±4.35	45.50±2.69
	Loss 25%	52.99±1.70	54.33±2.81	51.64±1.93
	Loss 75%	21.04±1.65	18.33±2.86	23.75±1.58

Percentages of selections for each symbol in the transfer task. Note that the transfer task data included only those trials in which choices were based on symbols from the same emotional condition during learning.
SEM: standard error of the mean

Table A.2.3. Effect of valence and anxiety on learning performance.

ANOVA	Combined	Experiment 1	Experiment 2
Anxiety	$F_{1,99} = 1.63$ $p = .2042$ $\eta_p^2 = 0.01$	$F_{1,49} = 0.91$ $p = .3454$ $\eta_p^2 = 0.01$	$F_{1,49} = 0.72$ $p = .4006$ $\eta_p^2 = 0.01$
Valence	$F_{1,99} = 0.00$ $p = .9657$ $\eta_p^2 = 0.00$	$F_{1,49} = 2.42$ $p = .1259$ $\eta_p^2 = 0.04$	$F_{1,49} = 2.18$ $p = .1459$ $\eta_p^2 = 0.04$
Anxiety × Valence	$F_{1,99} = 1.45$ $p = .2319$ $\eta_p^2 = 0.01$	$F_{1,49} = 0.12$ $p = .7339$ $\eta_p^2 = 0.00$	$F_{1,49} = 1.96$ $p = .1679$ $\eta_p^2 = 0.03$

Results from a 2x2 repeated measure ANOVA assessing the effects of anxiety and valence on correct response rate averaged across each condition of the learning task.

Table A.2.4. Effect of valence and anxiety on learning dynamics.

GLME: Fixed-Effect	Combined	Experiment 1	Experiment 2
Intercept	0.33±0.04 T_{25441}=9.26; p= 1.06× 10^{-14}	0.35±0.06 T_{9441}=5.31; p = 1.07× 10^{-7}	0.33±0.04 T_{15992}=7.73; p = 1.06× 10^{-14}
Trial	0.08±0.00 T_{25441}=12.23; p = 9.13× 10^{-23}	0.10±0.00 T_{9441}=11.86; p = 3.16× 10^{-32}	0.08±0.00 T_{15992}=9.83; p = 9.13× 10^{-23}
Valence	0.00±0.04 T_{25441}=0.30; p = .7632	0.02±0.06 T_{9441}=0.41; p = .6772	0.00±0.04 T_{15992}=0.02; p = .9806
Anxiety	0.01±0.06 T_{25441}=0.53; p = .5956	0.02±0.06 T_{9441}=0.33; p = .7409	0.01±0.06 T_{15992}=0.27; p = .7851
Trial × Valence	0.00±0.00 T_{25441}=1.66; p = .0961	0.03±0.00 T_{9441}=4.27; p = 1.89× 10^{-5}	0.00±0.00 T_{15992}=0.12; p = .8991
Trial × Anxiety	0.00±0.00 T_{25441}=0.89; p = .3732	0.00±0.00 T_{9441}=0.84; p = .3975	0.00±0.00 T_{15992}=0.68; p = .4962
Valence ×Anxiety	0.10±0.05 T_{25441}=0.60; p = .5426	-0.03±0.07 T_{9441}=-0.50; p = .6142	0.10±0.05 T_{15992}=1.84; p = .0646
Trial×Valence×Anxiety	0.00±0.00 T_{25441}=0.29; p = .7646	0.00±0.00 T_{9441}=0.72; p = .4713	0.00±0.00 T_{15992}=-0.16; p = .8705

GLME model result confirm the results from the ANOVA. Trial-by-trial choices (correct = 1, incorrect = 0) were entered into a logistic model with the fixed factors trial, valence and anxiety.

Table A.2.5. Effect of valence and anxiety on transfer performance.

ANOVA	Combined	Experiment 1	Experiment 2
Anxiety	$F_{1,99} = 0.00$ $p = 1.000$ $\eta_p^2 = 0.00$	$F_{1,49} = 0.00$ $p = 1.000$ $\eta_p^2 = 0.00$	$F_{1,49} = 0.00$ $p = 1.000$ $\eta_p^2 = 0.00$
Quality	$F_{1,99} = 241.60$ $p = 2.59\times10^{-28}$ $\eta_p^2 = 0.70$	$F_{1,49} = 126.68$ $p = 3.43\times10^{-15}$ $\eta_p^2 = 0.72$	$F_{1,49} = 163.98$ $p = 2.97\times10^{-17}$ $\eta_p^2 = 0.76$
Valence	$F_{1,99} = 212.85$ $p = 2.07\times10^{-26}$ $\eta_p^2 = 0.68$	$F_{1,49} = 95.28$ $p = 4.45\times10^{-13}$ $\eta_p^2 = 0.66$	$F_{1,49} = 121.40$ $p = 7.29\times10^{-15}$ $\eta_p^2 = 0.71$
Anxiety × Quality	$F_{1,99} = 6.37$ $p = .0132$ $\eta_p^2 = 0.06$	$F_{1,49} = 1.76$ $p = .1896$ $\eta_p^2 = 0.03$	$F_{1,49} = 6.30$ $p = .0154$ $\eta_p^2 = 0.11$
Anxiety × Valence	$F_{1,99} = 0.31$ $p = .5756$ $\eta_p^2 = 0.00$	$F_{1,49} = 0.32$ $p = .5742$ $\eta_p^2 = 0.00$	$F_{1,49} = 0.03$ $p = .8610$ $\eta_p^2 = 0.00$
Quality × Valence	$F_{1,99} = 2.16$ $p = .1439$ $\eta_p^2 = 0.02$	$F_{1,49} = 3.92$ $p = .0531$ $\eta_p^2 = 0.07$	$F_{1,49} = 0.29$ $p = .5867$ $\eta_p^2 = 0.00$
Quality×Valence×Anxiety	$F_{1,99} = 0.55$ $p = .4587$ $\eta_p^2 = 0.00$	$F_{1,49} = 0.11$ $p = .7331$ $\eta_p^2 = 0.00$	$F_{1,49} = 3.62$ $p = .0628$ $\eta_p^2 = 0.06$

Following Palminteri et al., (2015), cue preference rate was computed from the transfer task, and analyzed with a 3x2 repeated measure ANOVA. Note that the transfer task data included only those trials in which choices were based on symbols from the same emotional condition during learning.

Table A.2.6. Choices from transfer task were analyzed via GLME.

GLME: Fixed-Effect	Combined	Experiment 1	Experiment 2
'(Intercept)'	-0.01 ±0.03 T_{13592}=-0.47; p = .6372	0.03±0.08 T_{2392}=0.43; p = .6672	-0.02±0.03 T_{11192}=-0.79; p = .4273
Diff_Valence	1.52±0.17 T_{13592}=8.90; $p = 5.83\times 10^{-19}$	1.53 ±0.24 T_{2392}=6.15; $p = 8.96 \times 10^{-10}$	1.55 ±0.22 T_{11192}=6.78; $p = 1.22 \times 10^{-11}$
Diff_Quality	1.70 ±0.26 T_{13592}=6.53; $p = 6.73\times 10^{-11}$	2.28 ±0.33 T_{2392}= 6.90; $p = 6.33\times 10^{-12}$	1.34 ±0.20 T_{11192}=6.69; $p = 2.33 \times 10^{-11}$
Anxiety	-0.05 ±0.04 T_{13592}=-1.07; p = .2840	-0.02±0.12 T_{2392}=-0.21; p = .8269	-0.07±0.04 T_{11192}=-1.52; p = .1269
Diff_Valence× Diff_Quality	0.03±0.07 T_{13592}=0.44; p = .6594	0.18±0.16 T_{2392}=1.13; p = .2583	-0.03±0.06 T_{11192}=-0.52; p = .5983
Diff_Valence× Anxiety	0.04±0.16 T_{13592}=0.29; p = .7659	0.05±0.25 T_{2392}=0.20; p = .8384	0.03±0.22 T_{11192}=0.15; p = .8729
Diff_Quality × Anxiety	0.55±0.22 T_{13592}=2.44; p = .0144	0.53 ±0.43 T_{2392}=1.23; p = .2185	0.57±0.25 T_{11192}=2.23; p = .0256
Diff_Valence×Diff_Quality×Anxiety	-0.03 ±0.09 T_{13592}=-0.43; p = .6642	0.03±0.23 T_{2392}=0.14; p = .8867	-0.04 ±0.10 T_{11192}=-0.48; p = .6282

GLME (Generalized linear mixed model estimates) model result confirm the results from the ANOVA. Trial-by-trial choices (right = 1, left = 0) from the transfer task were entered into a logistic model with the fixed factors valence difference, quality difference and anxiety. Specifically, the model included independent variables accounting for differences between right and left cues, Diff_valence was defined as the difference in valence during learning and Diff _quality as the difference in the likelihood of avoiding a loss/attaining a gain during learning. The factor anxiety reflects whether cues were learned in the context of threat (threat = 1; safe =0). These variables were used to predict if the right cue was chosen. Diff: Difference between Right and Left stimuli.

Table A.2.7. Choice rate (focus on intermediate values) from transfer task.

ANOVA	Combined	Experiment 1	Experiment 2
Anxiety	$F_{1,99} = 0.55$ $p = .4586$ $\eta_p^2 = 0.00$	$F_{1,99} = 0.12$ $p = .7331$ $\eta_p^2 = 0.00$	$F_{1,99} = 3.63$ $p = .0628$ $\eta_p^2 = 0.06$
Valence	$F_{1,99} = 5.96$ $p = .0163$ $\eta_p^2 = 0.05$	$F_{1,99} = 6.66$ $p = .0129$ $\eta_p^2 = 0.11$	$F_{1,99} = 0.27$ $p = .6039$ $\eta_p^2 = 0.00$
Anxiety * Valence	$F_{1,99} = 2.63$ $p = .1077$ $\eta_p^2 = 0.02$	$F_{1,99} = 0.52$ $p = .4726$ $\eta_p^2 = 0.01$	$F_{1,99} = 3.27$ $p = .0766$ $\eta_p^2 = 0.06$

Following Palminteri et al., (2015), preferences of intermediate values from the transfer task were selectively analyzed with a 2x2 repeated measure ANOVA.

Table A.2.8. Paired t-test results for interaction of quality and valence.

Contract	Combined	Experiment 1	Experiment 2
Better >Worse (safe pairs) (SG75 +SL25)- (SG25 +SL75)	$t_{99} = 9.30$ $p < .001$ $d = 0.65$	$t_{49} = 6.94$ $p < .001$ $d = 0.69$	$t_{49} = 6.58$ $p < .001$ $d = 0.65$
Better >Worse (threat pairs) (AG75 +AL25)- (AG25 +AL75)	$t_{99} = 14.53$ $p < .001$ $d = 1.02$	$t_{49} = 9.70$ $p < .001$ $d = 0.97$	$t_{49} = 12.79$ $p < .001$ $d = 1.27$
[Better >Worse (threat pairs)] > [Better >Worse (safe pairs)]	$t_{99} = 2.41$ $p = .0165$ $d = 0.17$	$t_{49} = 1.27$ $p = .2058$ $d = 0.12$	$t_{49} = 2.42$ $p = .0171$ $d = 0.24$

Post-hoc tests characterizing the interaction between quality and valence in the transfer task. The results show that identification of better stimuli is more pronounced when these were learned in a threatening compared to a safe context. These results were found for experiment 1 and the combined data, with experiment 2 showing an effect in the same direction.
SG75: safe 75 % of gain; SG25: safe 25 % of gain; AG75: threat 75 % of gain; AG25: threat 25 % of gain; SL75: safe 75 % of loss; SL25: safe 25 % of loss; AL75: threat 75 % of loss; AL25: threat 25 % of loss.

Table A.2.9. Effect of incidental anxiety on reaction time.

Anxiety might generally impact performance by increasing cognitive demands, which is usually associated with longer reaction times. That is, cognitive demands are higher in the threat compared to the no-threat condition. We analyzed reaction time using ANOVA and post-hoc paired t-test.

A. ANOVA on Reaction time

ANOVA	Combined	Experiment 1	Experiment 2
Anxiety	$F_{1,99} = 2.66$ $p = .1056$ $\eta_p^2 = 0.02$	$F_{1,49} = 3.49$ $p = .0676$ $\eta_p^2 = 0.04$	$F_{1,49} = 0.56$ $p = .4569$ $\eta_p^2 = 0.01$
Valence	$F_{1,99} = 305.11$ $p = 5.30\times10^{-32}$ $\eta_p^2 = 0.75$	$F_{1,49} = 109.25$ $p = 4.5\times10^{-14}$ $\eta_p^2 = 0.58$	$F_{1,49} = 264.77$ $p = 2.1\times10^{-21}$ $\eta_p^2 = 0.84$
Anxiety × Valence	$F_{1,99} = 1.23$ $p = .2697$ $\eta_p^2 = 0.01$	$F_{1,49} = 0.17$ $p = .6818$ $\eta_p^2 = 0.00$	$F_{1,49} = 3.27$ $p = .0764$ $\eta_p^2 = 0.06$

Reaction time was averaged in each condition of the learning task, and analyzed with a 2x2 repeated measure ANOVA. The results showed a significant main effect of outcome valence. A non-significant main effect of anxiety and its interaction with outcome valence indicates anxiety has a limited effect on cognitive load

B. Paired t-test on Reaction time

Contrast	Combined	Experiment 1	Experiment 2
Threat > Safe	$t_{99} = -1.63$ $p = .1056$ $d = 0.16$	$t_{49} = -1.86$ $p = .0676$ $d = 0.25$	$t_{49} = -0.74$ $p = .4569$ $d = 0.11$

A post-hoc analysis with paired t-test showed a trend of faster reaction times in the threat compared to the safe condition, indicating that participants did not respond more slowly under threat. The absence of slower reaction times under conditions of anxiety, one of the hallmark measures of cognitive load (Pashler, 1994), indicates that participants did not face greater cognitive load in the threat condition.

Table A.2.10. Parameter comparison.

THREAT > SAFE	Combined	Experiment 1	Experiment 2
β	$t_{99} = 0.62$	$t_{49} = 0.90$	$t_{49} = 0.09$
	$p = .5309$	$p = .3683$	$p = .9249$
	$d = 0.06$	$d = 0.12$	$d = 0.01$
α^+	$t_{99} = -0.18$	$t_{49} = -0.03$	$t_{49} = -0.24$
	$p = .8537$	$p = .9749$	$p = .8072$
	$d = 0.01$	$d = 0.00$	$d = 0.03$
α^-	$t_{99} = 0.77$	$t_{49} = 0.31$	$t_{49} = 1.30$
	$p = .4412$	$p = .7534$	$p = .1991$
	$d = 0.07$	$d = 0.04$	$d = 0.18$
α_v	$t_{99} = -1.5$	$t_{49} = -1.63$	$t_{49} = -0.55$
	$p = .1362$	$p = .1078$	$p = .5833$
	$d = 0.15$	$d = 0.23$	$d = 0.07$
w	$t_{99} = 0.42$	$t_{49} = 1.09$	$t_{49} = -0.44$
	$p = .6698$	$p = .2787$	$p = .6590$
	$d = 0.04$	$d = 0.15$	$d = 0.06$

In the modelling approach (step2), each parameter from the wining model (i.e., RELa,w) was split into safe and threat and formed a new model (i.e., model2-6). To test which learning parameter was significantly altered by anxiety, we conducted paired t-test for each learning parameter to statistically compare threat and safe parameters. No significant effect of anxiety was found in neither combined nor separated dataset.

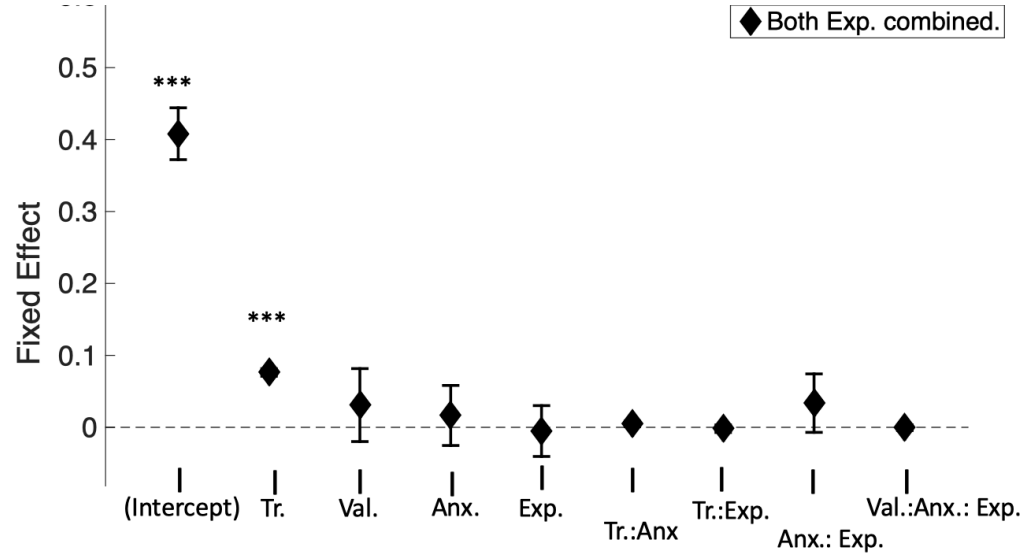
A.3 Robustness Tests: Effects of experiment

We assessed whether the two threat of shock paradigms we used caused systematic changes in learning performance, both for the learning and transfer tasks. To ensure different threat of shock paradigms did not bias our data, we conducted an additional GLME regressions by including Experiment as fixed effect in original factorial analyses ($GLME2_{Learning}$ and $GLME2_{Transfer}$). In terms of learning performance, a non-significant main effect of experiment was obtained, and the factor experiment did not interact with anxiety (**Figure A.3**). Similarly, the factor experiment did not show a significant effect on the performance in the transfer task, nor was there a significant interaction of experiment with anxiety (**Figure A.3**). These results indicate that the two different threat of shock paradigms used in experiments 1 and 2 have similar effects on learning and post-learning performance regardless of the dynamics and intensity of the anxiety induction that lead to different anxiety levels.

$GLME2_{Learning}$: Correct ~ (Intercept) + Valence + Anxiety * Experiment * Trial + (1 + Valence + Anxiety * Experiment * Trial | Subject);

$GLME2_{Transfer}$: ChooseRight ~ (Intercept) + Diff_Valence + Diff_Quality * Anxiety* Experiment + (1 + Diff_Valence + Diff_Quality * Anxiety* Experiment | Subject);

A. GLME results for Learning task



B. GLME results for Transfer task

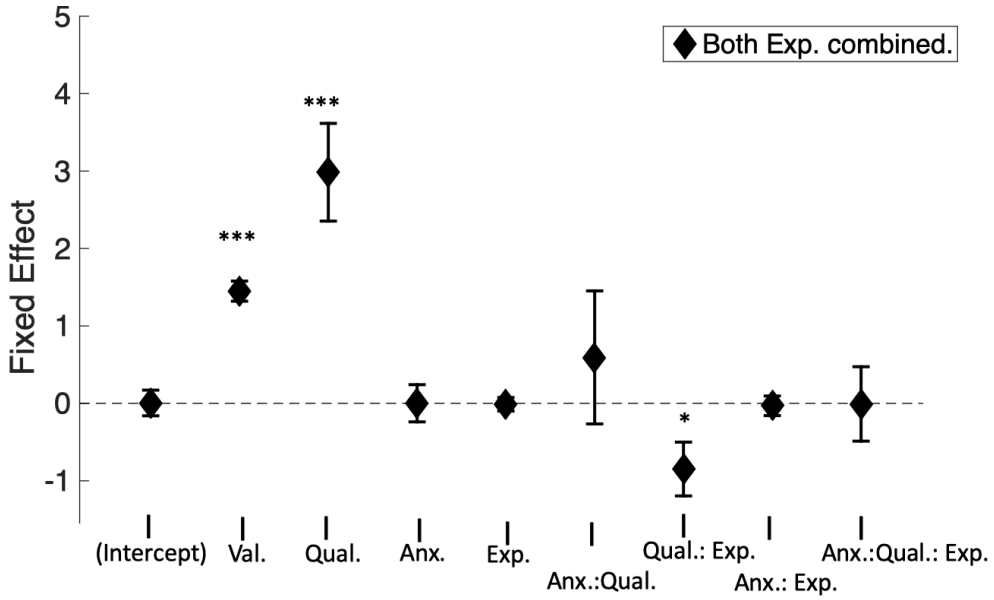


Figure A.3. The effect of experiment on learning **(A)** and preferences **(B)**. GLME model with correct choice (i.e., $GLME2_{Learning}$) and choosing right option (i.e., $GLME2_{Transfer}$) as dependent variables. The y-axis represents the estimated standardized coefficient (t-value), and the x-axis represents each factor in the GLME model.

Exp.: experimental number; **Val.**: outcome valence; **Qual.**: Quality of cue (i.e. Higher expected value in its pair during learning); **Anx.**: Anxiety manipulation.

~ .05< p <.1; ***** .01< p <.05; ****** .001< p < .01; ******* p < .001

A.4 Robustness tests: Effects of trait anxiety

Trait anxiety can modulate the impact of induced anxiety, as indicated in previous studies (Cavanagh et al., 2011). Specifically, highly anxious participants might react more strongly to a threatening environment than participants with low levels of anxiety. To assess these potentially interactive effects between trait and state anxiety, we assessed whether BAI moderates the effects of state anxiety or valence on learning and post-learning performance.

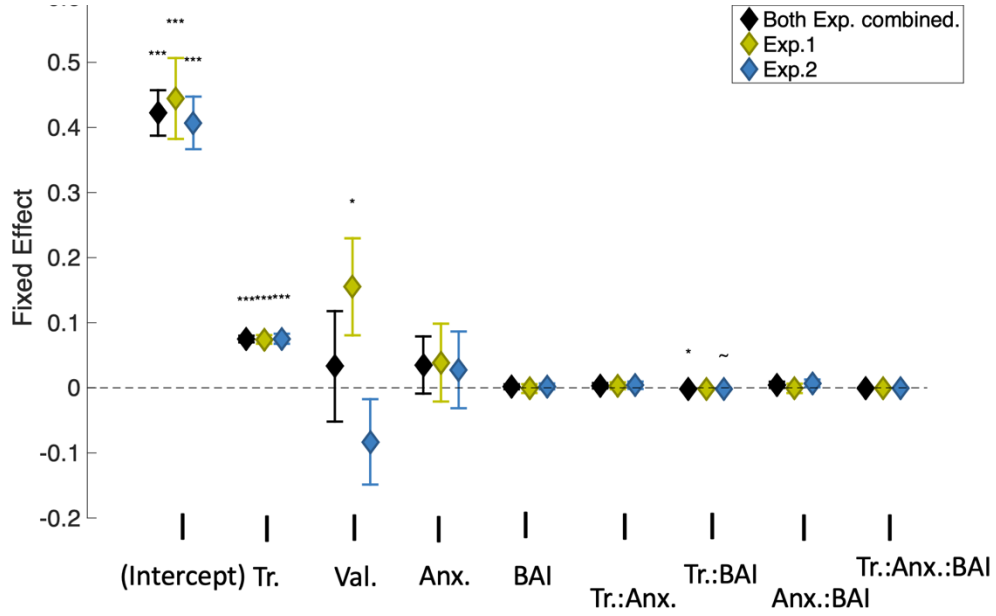
We ran the additional GLME regressions that included trait anxiety (i.e. BAI score) as a fixed factor for decisions in the learning and transfer tasks. In both models, BAI did not yield a significant main effect, nor were two- and three-way interactions with the factor anxiety significant. These results indicate that trait anxiety has limited effects on both learning and post-learning performances (**Figure A.4**) and does not modulate the impact of state anxiety on learning and preferences.

Interaction between Trait and State Anxiety

$GLME3_{Learning}$: Correct ~ (Intercept) + BAI * Anxiety * Trial + Valence + (1 + Anxiety * Trial + Valence | Experiment/Subject);

$GLME3_{Transfer}$: ChooseRight ~ (Intercept) + Anxiety * BAI + Diff_Valence * Diff_Quality + (1 + Diff_Valence * Diff_Quality + Anxiety | Experiment/Subject);

A. GLME results for Learning task



B. GLME results for Transfer task

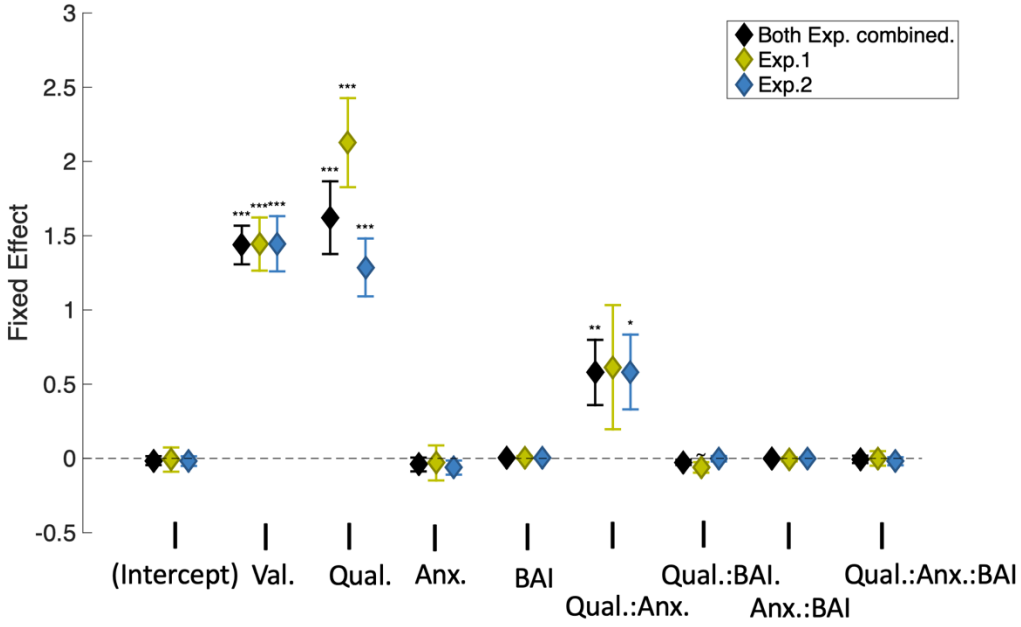


Figure A.4. The effect of trait anxiety on learning **(A)** and preferences **(B).** Generalized linear mixed-effect model with correct choice (i.e., $GLME3_{Learning}$) and choosing right option (i.e., $GLME3_{Transfer}$) as dependent variable. The y-axis represents the estimated standardized coefficient, and x-axis represents each factor in the GLME model. Each dot was plotted with estimated standardized coefficient and its standard error. **BAI**: Beck Anxiety inventory score; **Val.**: outcome valence; **Qual.**: Quality of cue (i.e. Higher expected value in its pair during learning); **Anx.**: Anxiety manipulation.
~ $.05< p <.1$; * $.01< p <.05$; ** $.001< p < .01$; *** $p < .001$

A.5 Robustness Tests: Assessment of the Cognitive Resources Hypothesis

We tested an alternative interpretation for our results that states that the non-significant effects of anxiety on learning may be due to a relatively easy task that does not tax our participants' cognitive resources (Eysenck et al., 2007). Although we did not directly manipulate cognitive demands in the current experiment, we can inspect individual differences in average task performance as a proxy for how demanding the learning task was on average for individual participants.

To test this interpretation, we first used cluster analysis (k-means, 2 clusters) to identify good and bad performers based on participants' average learning performance throughout the task independent of condition. The cluster centroids for good and bad performers were found by minimizing the sum of data-to-centroid distances (Lloyd, 1982). This way, the mean performance is differentiable between clusters (Good performers: 75.51±0.7%; Bad performers: 60.34±1.0%). We then ran additional GLME regressions that included a dummy for average performance to predict correct choice in the learning task and preference in the transfer task. For the learning data, the results show that the performance group dummy can predict correct choice and the effect is positively modulated by trial, which confirms the clustering and furthermore implies that better average performance is also associated with faster learning. Unlike predictions from attentional control theory, the performance group effect is not modulated by anxiety (two-way and three-way interactions of anxiety and cluster and trial are insignificant) (**Figure A.5**).

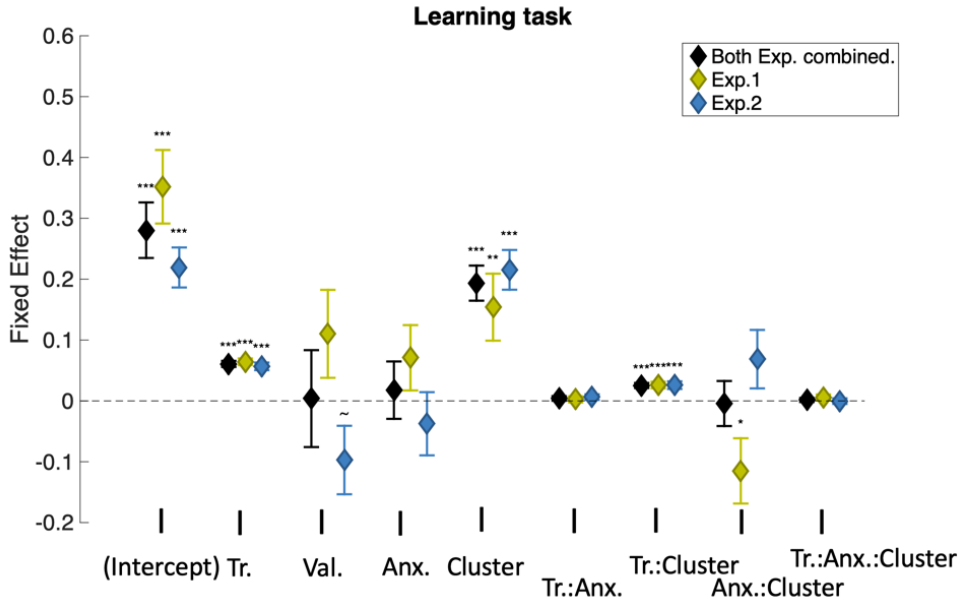
Results for post-learning data show non-significant main and interaction effects of performance group on predicting choice. Importantly, the non-significant two- and three-way interactions of performance group and quality (and anxiety) imply that better and worse learners have similar abilities to recognizing the better symbol, and anxiety influences these groups equally (**Figure A.5**).

Interaction between Trait and State Anxiety

$GLME4_{Learning}$: Correct ~ (Intercept) + cluster * Anxiety * Trial + Valence + (1 + Anxiety * Trial + Valence | Experiment/Subject);

$GLME4_{Transfer}$: ChooseRight ~ (Intercept) + Anxiety * cluster + Diff_Valence * Diff_Quality + (1 + Diff_Valence * Diff_Quality + Anxiety | Experiment/Subject);

A. GLME results for Learning task



B. GLME results for Transfer task

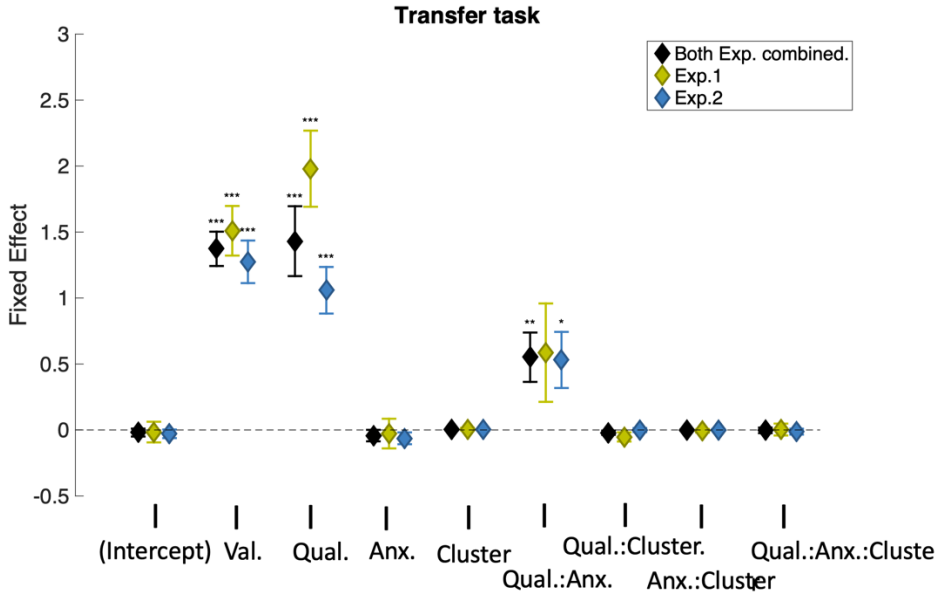


Figure A.5. The effect of trait anxiety on learning **(A)** and preferences **(B)**. Generalized linear mixed-effect model with correct choice (i.e., $GLME4_{Learning}$) and choosing right option (i.e., $GLME4_{Transfer}$) as dependent variable. The y-axis represents the estimated standardized coefficient, and x-axis represents each factor in the GLME model. Each dot was plotted with estimated standardized coefficient and its standard error. **Cluster**: Good or bad learner; **Val.**: outcome valence; **Qual.**: Quality of cue (i.e. Higher expected value in its pair during learning); **Anx.**: Anxiety manipulation.
~ .05< p <.1; ***** .01< p <.05; ****** .001< p < .01; ******* p < .001

A.6 Robustness Tests: Selectively test high SCR-responders

Our participants might respond differently to our anxiety induction, which may limit the behavioral effects of threat of shock in some participants. To address this, we further selectively analyzed participants that showed elevated SCR responses during the threat condition (threat > safe). We first performed a median split on our participants and analyzed the learning data from only those subject that fell in the top 50th percentile of the SCR response in each experiment. Using our original analysis approach outlined in this chapter (averaged performance entered into ANOVA, trial-by-trial data entered into GLME), we then investigate the effects of anxiety in the top half of SCR responders. Even in this select subsample that would be expected to show the greatest effect of anxiety (Berghorst et al., 2013), the effect of anxiety on learning and post-learning performance and its interactions remain non-significant as shown in the tables and figures below.

Table A.6.1. ANOVA results for high SCR responders in the learning task.

	Combined	**Experiment 1**	**Experiment 2**
Anxiety	$F_{1,45} = 0.00$ $p = .9365$ $\eta_p^2 = 0.00$	$F_{1,23} = 1.23$ $p = .2776$ $\eta_p^2 = 0.05$	$F_{1,21} = 0.55$ $p = .4679$ $\eta_p^2 = 0.02$
Valence	$F_{1,45} = 0.03$ $p = .8536$ $\eta_p^2 = 0.00$	$F_{1,23} = 0.36$ $p = .5538$ $\eta_p^2 = 0.01$	$F_{1,21} = 0.28$ $p = .6044$ $\eta_p^2 = 0.01$
Anxiety × Valence	$F_{1,45} = 0.01$ $p = .9160$ $\eta_p^2 = 0.00$	$F_{1,23} = 0.83$ $p = .3693$ $\eta_p^2 = 0.03$	$F_{1,21} = 0.96$ $p = .3375$ $\eta_p^2 = 0.04$

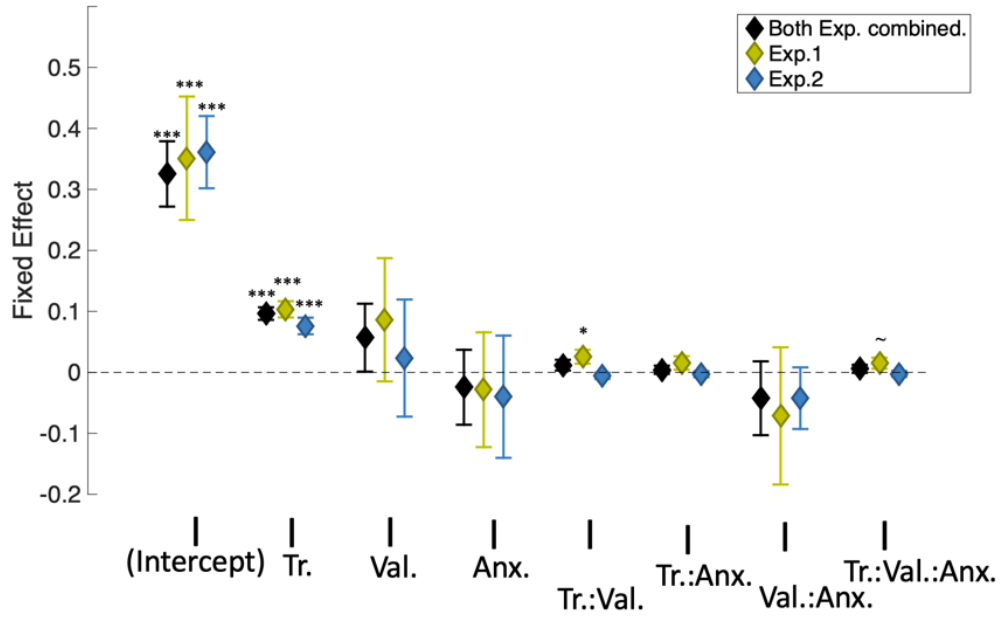
Using the same 2x2 repeated measure ANOVA to assess the effects of anxiety and valence on correct response rate averaged across each condition of the learning task.

Table A.6.2. ANOVA results for high SCR responders in the transfer task.

	Combined	**Experiment 1**	**Experiment 2**
Anxiety	$F_{1,45} = 0.00$ $p = 1.000$ $\eta_p^2 = 0.00$	$F_{1,23} = 0.00$ $p = 1.000$ $\eta_p^2 = 0.00$	$F_{1,21} = 0.00$ $p = 1.000$ $\eta_p^2 = 0.00$
Quality	$F_{1,45} = 85.11$ $p = 6.09 \times 10^{-12}$ $\eta_p^2 = 0.65$	$F_{1,23} = 46.31$ $p = 6.10 \times 10^{-7}$ $\eta_p^2 = 0.66$	$F_{1,21} = 58.37$ $p = 1.70 \times 10^{-7}$ $\eta_p^2 = 0.73$
Valence	$F_{1,45} = 84.38$ $p = 6.93 \times 10^{-12}$ $\eta_p^2 = 0.65$	$F_{1,23} = 35.47$ $P = 4.50 \times 10^{-7}$ $\eta_p^2 = 0.60$	$F_{1,21} = 51.57$ $p = 4.43 \times 10^{-7}$ $\eta_p^2 = 0.71$
Anxiety × Quality	$F_{1,45} = 1.60$ $p = .2116$ $\eta_p^2 = 0.03$	$F_{1,23} = 0.46$ $p = .5036$ $\eta_p^2 = 0.01$	$F_{1,21} = 1.79$ $p = .1950$ $\eta_p^2 = 0.07$
Anxiety × Valence	$F_{1,45} = 0.30$ $P = .5810$ $\eta_p^2 = 0.00$	$F_{1,23} = 0.02$ $p = .8852$ $\eta_p^2 = 0.00$	$F_{1,21} = 1.11$ $p = .3028$ $\eta_p^2 = 0.05$
Quality × Valence	$F_{1,45} = 3.64$ $p = .0628$ $\eta_p^2 = 0.07$	$F_{1,23} = 3.19$ $p = .0872$ $\eta_p^2 = 0.12$	$F_{1,21} = 0.50$ $p = .4844$ $\eta_p^2 = 0.02$
Quality×Valence×Anxiety	$F_{1,45} = 0.14$ $p = .7035$ $\eta_p^2 = 0.00$	$F_{1,23} = 0.21$ $p = .6446$ $\eta_p^2 = 0.00$	$F_{1,21} = 0.01$ $p = .9134$ $\eta_p^2 = 0.00$

Using the same 2x3 repeated measure ANOVA to assess the effects of anxiety, valence and quality on cue preference (i.e., percentage of selection) in the transfer task. Note that the transfer task data included only those trials in which choices were based on symbols from the same emotional condition during learning.

A. GLME results for learning task



B. GLME results for transfer task

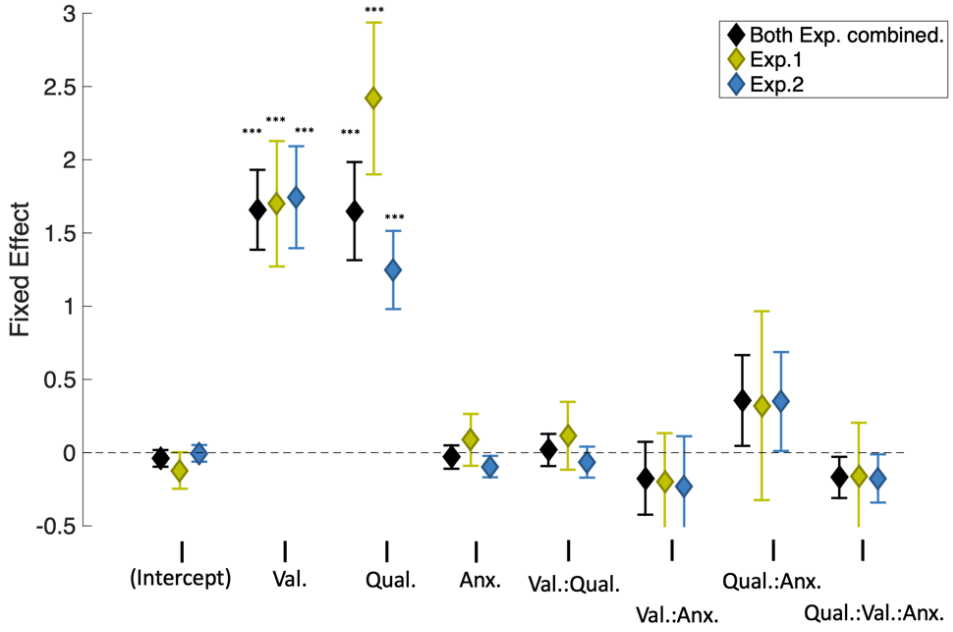


Figure A.6. Effect of incidental anxiety on learning performance from high SCR-responders. (A) Generalized linear mixed-effect model (i.e., $GLME1_{Learning}$) for the learning task (DV = correct/incorrect choice). The y-axis represents the estimated standardized coefficient, and x-axis represents each factor in the GLME model. **(B)** Generalized linear mixed-effect model (i.e., $GLME1_{Transfer}$) for the transfer task (DV = choosing the right option). The y-axis represents the estimated standardized coefficient (t-value), and x-axis represents each factor in the GLME model. Each dot was plotted with estimated standardized coefficient and its standard error. Each dot was plotted with estimated standardized coefficient and its standard error.

~ $.05 < p < .1$; ***** $.01 < p < .05$; ****** $.001 < p < .01$; ******* $p < .001$

A.7 Robustness Tests: Performance in early and late learning stages

The significant effect of anxiety on post-learning performance in the absence of an effect of anxiety on average learning performance poses the question to what extent the effect is related to learning performance. We first hypothesized that there is a difference between early learning performance and late learning performance, with late learning performance driving the post-learning preferences. To capture late learning performance, we focus specifically on the last session(s) (experiment 1: last session with 96 trials, experiment 2: last 2 sessions with 160 trials), because this period reflects the most recent learning performance relative to the transfer task (i.e., the stimuli presented in the post-learning task are those learned in these final session(s)). We selectively analyzed the learning data from the final session(s) only (session 2 from experiment 1 and sessions 3-4 from experiment 2). Entering this restricted dataset into an ANOVA identifies a marginally significant main effect of anxiety ($F_{1,99}$ = 3.43, p = .0668, η_p^2 = 0.03) and a near-significant interaction between anxiety and outcome valence ($F_{1,99}$ = 3.93, p = .0501, η_p^2 = 0.01). Follow-up paired t-tests showed that the discrepancy between G75 and L25 (i.e., better options in the gain and loss domain, respectively) was significantly higher in safe than in threat conditions (t_{99} = 1.98, p = .0501, d = 0.19), and the effect was driven by gain learning (Safe G75> Threat G75: t_{99} = -2.03, p = .0443, d = 0.20) but not by loss learning (Safe L25> Threat L25: t_{99} = 0.13, p = .8930, d = 0.01). Moreover, the result was not found when we selectively analyzed data from early learning stages.

The results indicate that the learning performance in the late learning stage, rather than early or overall performance, might be more susceptible to the effect of incidental anxiety. On the other hand, in the final session, incidental anxiety boosted gain learning as G25 was selected relatively less and G75 relatively more in the threat condition compared to the safe condition, while no effect of anxiety was observed for loss learning.

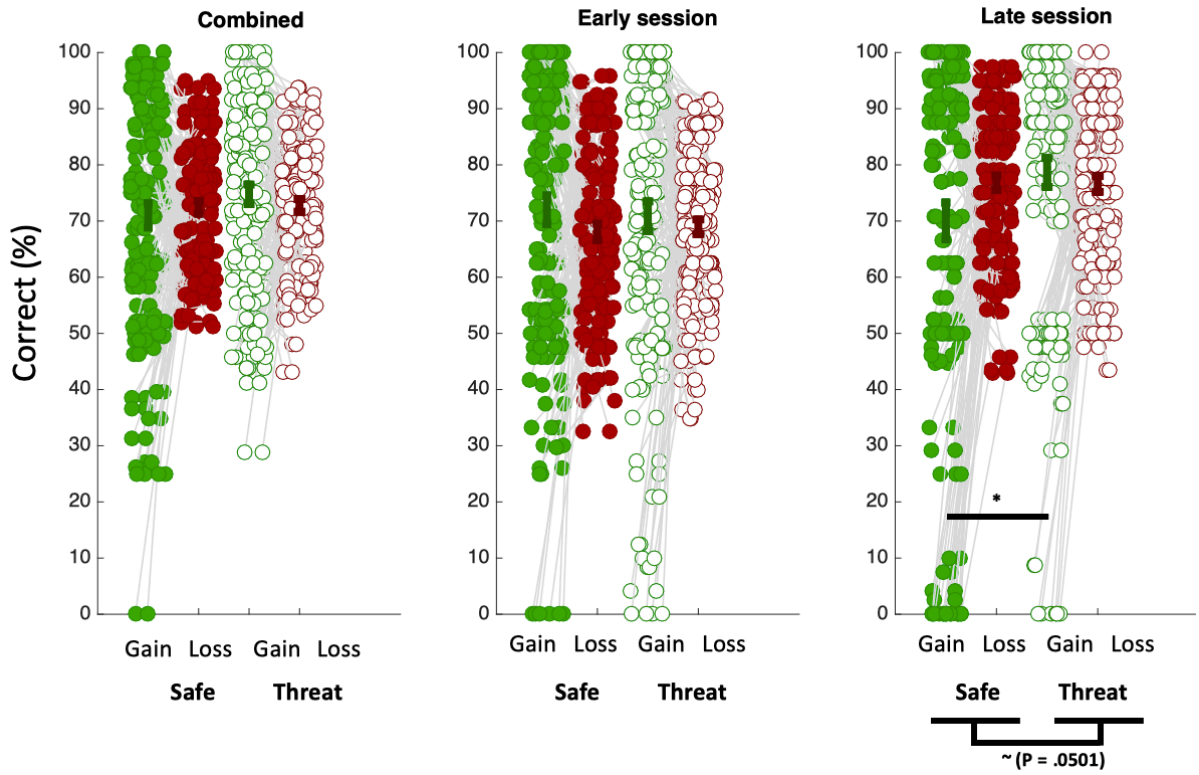


Figure A.7. Learning performance at different stages of learning task. Average correct rate across four conditions for combined data (left panel), early learning session (middle panel) and late session (right panel).

Table A.7. Learning performance at different stages of learning task.

	Combined	**Early session**	**Late session**
Anxiety	$F_{1,99} = 1.63$ $p = .2042$ $\eta_p^2 = 0.01$	$F_{1,99} = 0.00$ $p = .9539$ $\eta_p^2 = 0.00$	$F_{1,99} = 3.43$ $p = .0668$ $\eta_p^2 = 0.03$
Valence	$F_{1,99} = 0.00$ $p = .9657$ $\eta_p^2 = 0.00$	$F_{1,99} = 1.57$ $p = .2130$ $\eta_p^2 = 0.01$	$F_{1,99} = 1.07$ $p = .3015$ $\eta_p^2 = 0.01$
Anxiety × Valence	$F_{1,99} = 1.45$ $p = .2319$ $\eta_p^2 = 0.01$	$F_{1,99} = 0.32$ $p = .5733$ $\eta_p^2 = 0.00$	$F_{1,99} = 3.93$ $p = .0501$ $\eta_p^2 = 0.03$

Correct response rate was averaged in each condition of the learning task, and analyzed with a 2x2 repeated measure ANOVA.

A.8 Robustness tests: Association between 'late' learning stage and preference

The number of selections in the learning stage might be further transferred to preferences in the transfer task. We therefore test whether the number of selections of high value stimuli (relative to lower value stimuli) in the final sessions can predict preferences of G75 (vs. G25) and L25 (vs. L75) in the transfer task. To test the effects of learning on post-learning preferences (%), we ran a GLME with learning (i.e., the number of selections in the final sessions), valence and anxiety as predictors.

$$GLME5_{Transfer}: \text{Preference} \sim (\text{Intercept}) + \text{Valence} * \text{Learning} * \text{Anxiety} + (1 + \text{Valence} * \text{Learning} * \text{Anxiety} \mid \text{Experiment/Subject});$$

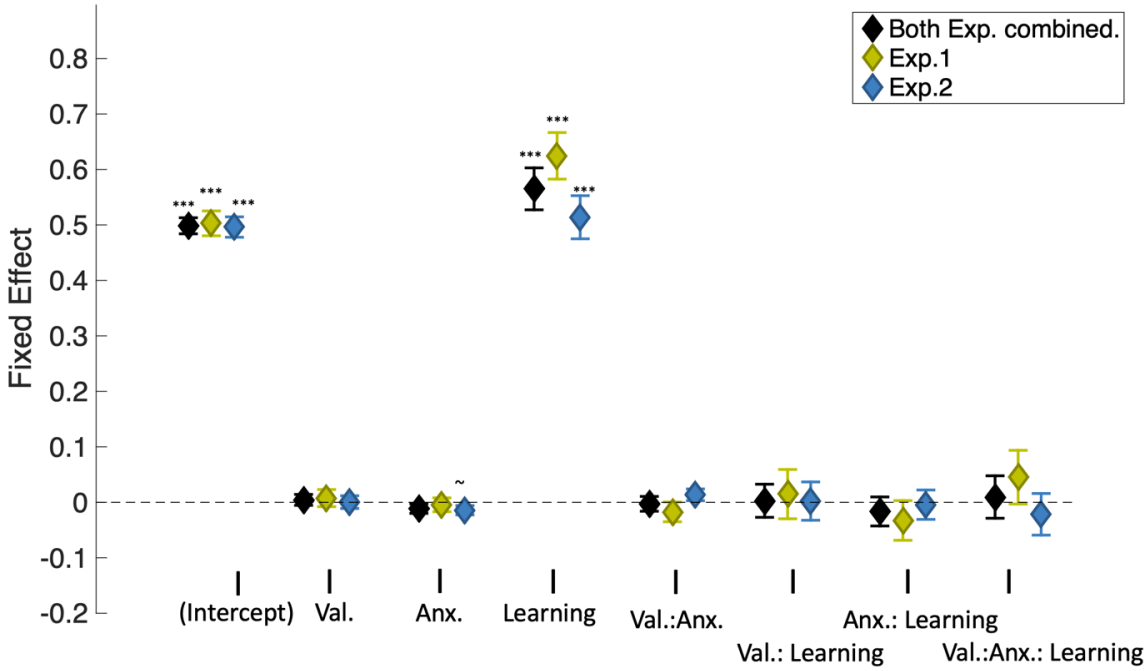
The results showed a significant main effect of learning ($\beta_{Learning} = 0.59 \pm 0.03$, $t_{392} = 19.61$, $p = 3.65 \times 10^{-60}$), but no significant interactions with outcome valence ($\beta_{Learning \times Valence} = 0.02 \pm 0.02$, $t_{392} = 1.25$, $p = .2102$) nor anxiety ($\beta_{Learning \times Anxiety} = 0.01 \pm 0.02$, $t_{392} = 0.56$, $p = .5722$).

We also analyzed differential preferences between intermediate values (i.e., GL_Pref: G25> L25) reflective of context effects via an additional GLME, which included differential number of selections in the later learning stage (i.e., GL_Learning: G25>L25) and anxiety as factors.

$$GLME6_{Transfer}: \text{GL\_Pref} \sim (\text{Intercept}) + \text{GL\_Learning} * \text{Anxiety} + (1 + \text{GL\_Learning} * \text{Anxiety} \mid \text{Experiment/Subject});$$

The results showed that the preferences between intermediate values can be generally predicted by learning performances (main effect of differences on learning performances: $\beta_{GL\_Learning} = 0.53 \pm 0.05$, $t_{196} = 9.31$, $p = 2.50 \times 10^{-17}$) regardless of the anxiety manipulation ($\beta_{Anxiety \times GL\_Learning} = 0.01 \pm 0.05$, $t_{196} = 0.31$, $p = .7550$). Jointly, these results showed that (1) there is a marginally significant effect of anxiety on late learning performance in the absence of an effect of anxiety on early learning performance, (2) there is a strong relationship between late learning performance and post-learning preferences, and (3) incidental anxiety might have a limited effect on changing the association between learning performance and post-learning preferences.

A. GLME on overall preferences



B. GLME on intermediate values

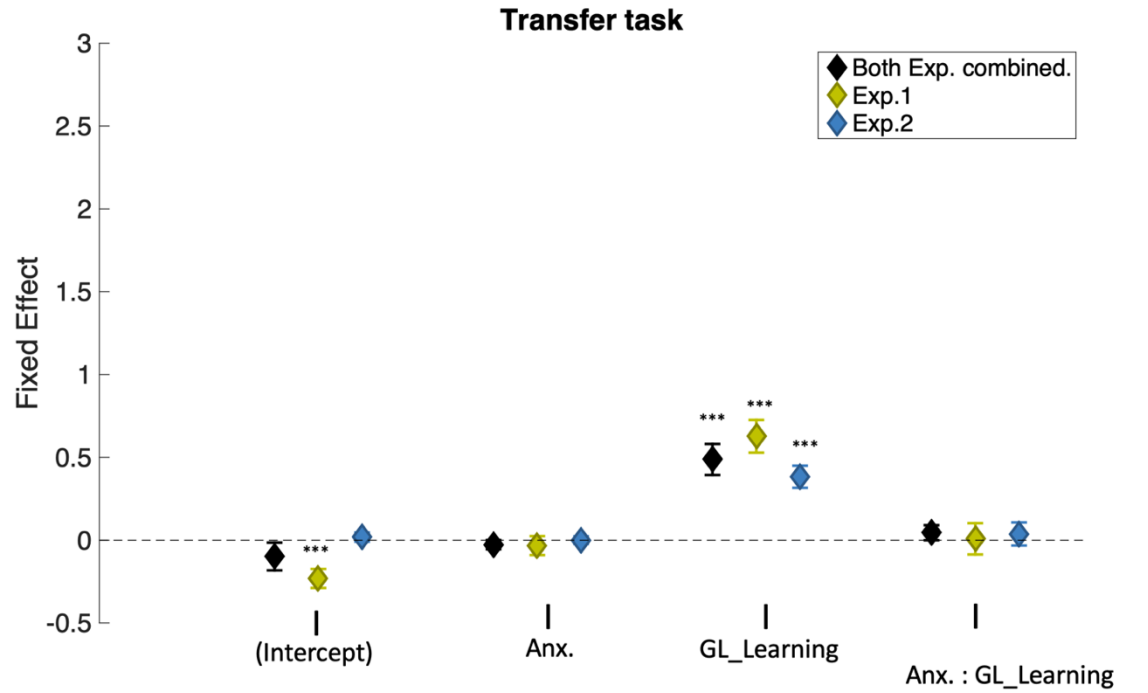


Figure A.8. The effect of number of selections on preferences. (A) Generalized linear mixed-effect model (i.e., $GLME5_{Transfer}$) with choosing rate as dependent variable. **(B)** Generalized linear mixed-effect model (i.e., $GLME6_{Transfer}$) with differential preferences between intermediate values (i.e., LG_pref: L25>G25) as dependent variable. The y-axis represents the estimated standardized coefficient, and the x-axis represents each factor in the GLME model. Dot colors indicate results from different dataset. **GL_Learning.**: learning performance (differential number of selections between intermediate values: G25> L25); **Anx.**: Anxiety manipulation.
~ .05< p <.1; ***** .01< p <.05; ****** .001< p < .01; ******* p < .001

APPENDIX B: Supplementary Material for Chapter 3

B.1 Overview of Experimental paradigms

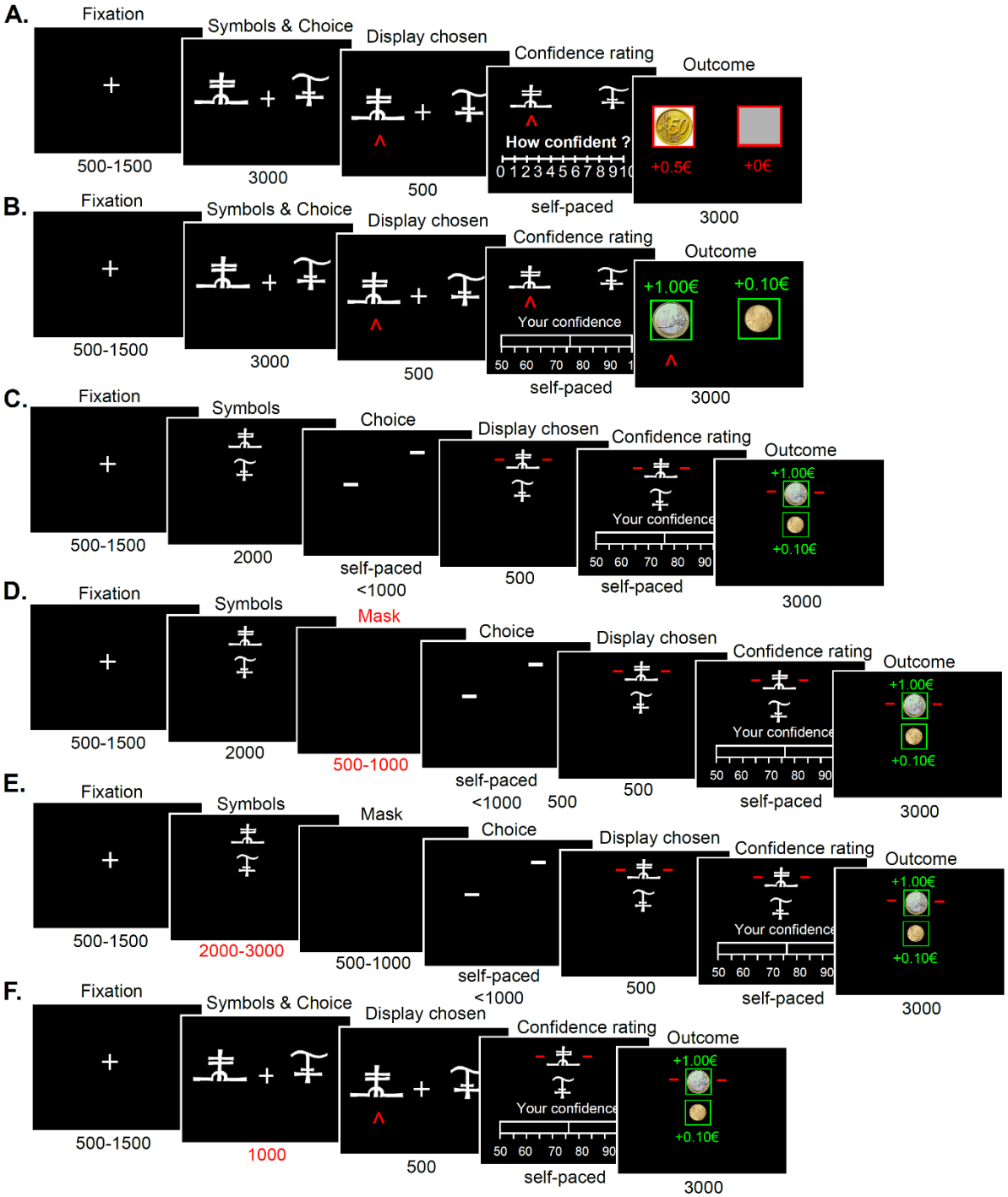


Figure B.1. Experimental paradigms. (A-F) Behavioral tasks for Experiments 1-6. Successive screens displayed in one trial are shown from left to right with durations in ms. All tasks are based on the same principle: after a fixation cross, participants are presented with a couple of abstract symbols displayed on a computer screen and have to choose between them. They are thereafter asked to report their confidence in their choice on a numerical scale. Outcome associated with the chosen symbol is revealed, sometimes paired with the outcome associated with the unchosen symbol -depending on the condition. Tasks specificities are as follow: **(A)** Experiment 1: symbols are displayed on the left and right sides of the screen. Confidence is reported on a 0-10 Likert scale non-incentivized. **(B)** Experiment 2: similar to experiment 1, except that confidence is reported on a 50-100% rating scale and incentivized. **(C)** Experiment 3: similar to Experiment 2, except that options are displayed on a vertical axis. Besides, the response mapping (how the left vs right arrow map to the upper vs lower symbol) is only presented after the symbol display, and the response has to be given within one second of the response mapping screen onset. **(D)** Experiment 4: similar to experiment 3, except that a short empty screen is used as a mask, between the symbol display and the response mapping. **(E)** Experiment 5: similar to experiment 4, except that a jitter is introduced in the symbol presentation. **(F)** Experiment 6: similar to experiment 2, except that a shorter duration is allowed from the symbol presentation to the choice.

B.2 Overview of treatment effect

Table B.2. One-way ANOVA results for the combined data from six experiments.

		Gain - Loss	Partial - Complete	Interaction
Performance (Accuracy rate)	F(5, 102), [η^2] (*p*-val)	0.35, [0.02] (0.884)	0.52, [0.03] (0.729)	0.56, [0.03] (0.725)
Confidence	F(5, 102), [η^2] (*p*-val)	1.26, [0.06] (0.289)	1.19, [0.06] (0.319)	0.51, [0.02] (0.771)
RT	F(5, 102), [η^2] (*p*-val)	7.98, [0.28] (<.001) ***	2.81, [0.12] (0.021)*	2.98, [0.13] (0.015)*

The role of experimental manipulations on the effect of outcome valence on Performance, Confidence and RT.

~ p <.1; * p <.05; ** p <.01; *** p <.001

B.3 Results of GLME

Table B.3.1. Estimated coefficients from generalized linear mixed-effect models on performance across experiments

		Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5	Exp. 6
Val.	β ± SE	0.40 ± 0.21	0.08 ± 0.18	0.16 ± 0.27	0.15 ± 0.19	-0.08 ± 0.17	0.12 ± 0.15
	t-val	1.86	0.32	0.58	0.78	-0.50	0.76
	(*p*-val)	(0.063)~	(0.748)	(0.561)	(0.43)	(0.620)	(0.449)
Inf.	β ± SE	0.31 ± 0.18	0.72 ± 0.16	0.52 ± 0.22	0.63 ± 0.30	0.51 ± 0.18	0.22 ± 0.16
	t-val	1.74	4.59	2.40	2.10	2.92	1.41
	(*p*-val)	(0.081)~	(<.001)***	(0.016)*	(0.036)*	(0.004)**	(0.158)
Val x Inf	β ± SE	0.10 ± 0.20	0.20 ± 0.21	0.16 ± 0.19	0.36 ± 0.23	0.04 ± 0.18	0.25 ± 0.12
	t-val	0.47	0.92	0.83	1.56	0.23	2.13
	(*p*-val)	(0.638)	(0.356)	(0.405)	(0.118)	(0.814)	(0.034)*
Fix (s)	β ± SE				0.18 ± 0.35	-0.28 ± 0.28	
	t-val	-	-	-	-0.51	-0.99	-
	(*p*-val)				(0.611)	(0.322)	
Stim (s)	β ± SE					0.03 ± 0.07	
	t-val	-	-	-	-	0.37	-
	(*p*-val)					(0.713)	
Mask (s)	β ± SE				0.14 ± 0.29	0.17 ± 0.28	
	t-val	-	-	-	-0.47	0.57	-
	(*p*-val)				(0.637)	(0.567)	
Sess.	β ± SE	0.34 ± 0.14	0.78 ± 0.15	0.58 ± 0.14	0.46 ± 0.15	0.30 ± 0.10	0.09 ± 0.07
	t-val	2.40	5.08	4.17	3.00	2.89	1.32
	(*p*-val)	(0.016)*	(<.001)***	(<.001)***	(0.003)**	(0.004)**	(0.186)

β: estimated regression coefficients for fixed effects. SE: estimated standard error of the regression coefficients.

Val.: valence; **Inf.**: information; **Fix.**: fixation duration; **Stim.**; stimulus display duration; **Sess**.: session number.

~ p <.1; * p <.05; ** p <.01; *** p <.001

Table B.3.2. Estimated coefficients from generalized linear mixed-effect models on confidence across experiments

		Exp. 1	**Exp. 2**	**Exp. 3**	**Exp. 4**	**Exp. 5**	**Exp. 6**
Val.	β ± SE	8.85 ± 1.51	8.29 ± 2.05	4.23 ± 1.17	5.34 ± 1.19	7.19 ± 2.27	7.05 ± 1.32
	t-val	5.86	4.04	3.59	4.50	3.16	5.34
	(*p*-val)	(<.001)***	(<.001)***	(<.001)***	(<.001)***	(0.002)**	(<.0001) ***
Inf.	β ± SE	0.76 ± 0.51	2.75 ± 1.20	0.95 ± 0.61	1.55 ± 0.85	1.59 ± 0.59	0.27 ± 0.78
	t-val	1.49	2.28	1.55	1.82	2.69	0.35
	(*p*-val)	(0.135)	(0.022)*	(0.120)	(0.069)~	(<0.001)***	(0.726)
Val x Inf	β ± SE	-2.16 ± 0.76	-1.38 ± 0.65	-0.90 ± 0.69	-1.51 ± 0.72	-2.67 ± 0.95	-2.05 ± 0.88
	t-val	-2.85	-2.12	-1.31	-2.10	-2.81	-2.33
	(*p*-val)	(0.004)**	(0.034)*	(0.192)	(0.036)*	(0.004)**	(0.019)*
Fix (s)	β ± SE				-1.11 ± 1.40	0.49 ± 1.43	
	t-val	-	-	-	-0.79	0.34	-
	(*p*-val)				(0.428)	(0.734)	
Stim (s)	β ± SE					0.21 ± 0.39	
	t-val	-	-	-	-	0.53	-
	(*p*-val)					(0.596)	
Mask (s)	β ± SE				0.27 ± 1.48	-0.40 ± 1.32	
	t-val	-	-	-	-0.18	-0.30	-
	(*p*-val)				(0.854)	(0.761)	
Sess.	β ± SE	2.99 ± 0.98	2.84 ± 0.68	1.75 ± 0.73	1.96 ± 0.89	1.20 ± 0.80	-0.49 ± 1.10
	t-val	3.05	4.19	2.41	2.23	1.50	-0.45
	(*p*-val)	(0.002)**	(<.001)***	(0.016)*	(0.026)*	(0.133)	(0.653)

β: estimated regression coefficients for fixed effects. SE: estimated standard error of the regression coefficients.

Val.: valence; **Inf.**: information; **Fix.**: fixation duration; **Stim.**; stimulus display duration; **Sess**.: session number.

~ p <.1; * p <.05; ** p <.01; *** p <.001

Table B.3.3. Estimated coefficients from generalized linear mixed-effect models on response times across experiments

		Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5	Exp. 6
Val.	β ± SE	-151.12 ± 40.37	-115.63 ± 30.96	-15.31 ± 4.33	-13.49 ± 4.97	-3.23 ± 3.44	-35.19 ± 8.60
	t-val	-3.74	-3.73	-3.53	-2.71	-0.94	-4.10
	(*p*-val)	(<.001)***	(<.001)***	(<.001)***	(0.007)**	(0.349)	(<.001)***
Inf.	β ± SE	-6.57 ± 19.58	-44.37 ± 15.75	5.81 ± 4.13	-2.81 ± 4.28	0.80 ± 3.78	-8.23 ± 5.11
	t-val	-0.34	-2.82	1.41	-0.65	-0.21	-1.61
	(*p*-val)	(0.737)	(0.005)**	(0.160)	(0513)	(0.832)	(0.107)
Val x Inf	β ± SE	65.58 ± 28.77	10.59 ± 18.88	3.75 ± 3.25	3.67 ± 4.19	-3.04 ± 3.85	8.35 ± 5.43
	t-val	2.28	0.56	1.15	0.88	-0.79	1.54
	(*p*-val)	(0.023)*	(0.575)	(0.249)	(0.381)	(0.430)	(0.124)
Fix (s)	β ± SE				-2.37 ± 16.13	18.56 ± 12.77	
	t-val	-	-	-	-0.15	1.45	-
	(*p*-val)				(0.883)	(0.146)	
Stim (s)	β ± SE					-13.12 ± 4.00	
	t-val	-	-	-	-	-3.28	-
	(*p*-val)					(<.001)***	
Mask(s)	β ± SE				-68.34 ± 16.35	-54.20 ± 14.5	
	t-val	-	-	-	-4.18	-3.73	-
	(*p*-val)				(<.001)***	(<.001)***	
Sess.	β ± SE	-152.43 ± 33.63	-146.28 ± 26.13	-26.93 ± 6.14	-32.55 ± 9.51	-27.57 ± 5.64	-6.39 ± 8.34
	t-val	-4.53	-5.60	-4.38	-3.42	-4.79	-0.77
	(*p*-val)	(<.001)***	(<.001)***	(<.001)***	(<.001)***	(<.001)***	(0.443)

β: estimated regression coefficients for fixed effects. SE: estimated standard error of the regression coefficients.

Val.: valence; **Inf.**: information; **Fix.**: fixation duration; **Stim.**; stimulus display duration; **Sess**.: session number.

~ p <.1; * p <.05; ** p <.01; *** p <.001

Table B.3.4. Estimated coefficients from generalized linear mixed-effect models on confidence, controlling for reaction times, across experiments

		Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5	Exp. 6
Val.	β ± SE	158.94 ± 28.40	7.12 ± 1.97	3.99 ± 1.18	5.16 ± 1.17	7.11 ± 1.70	6.29 ± 1.28
	t-val	5.60	3.61	3.39	4.41	4.19	4.91
	(*p*-val)	(<.001)***	(<.001)***	(<.001)***	(<.001)***	(<.001)***	(<.001)***
Inf.	β ± SE	10.47 ± 9.22	2.29 ± 1.15	0.99 ± 0.58	1.46 ± 0.84	1.67 ± 0.51	0.13 ± 0.82
	t-val	1.14	1.99	1.73	1.74	3.31	0.16
	(*p*-val)	(0.256)	(0.046)*	(0.084)~	(0.081)~	(<.001)***	(0.875)
Val x Inf	β ± SE	-31.32 ± 13.3	-1.09 ± 0.63	-0.79 ± 0.71	-1.37 ± 0.73	2.45 ± 0.81	-2.06 ± 0.88
	t-val	-2.35	-1.74	-1.11	-1.86	-3.02	-2.34
	(*p*-val)	(0.019)*	(0.082)~	(0.27)	(0.062)~	(0.003)**	(0.019)*
RT	β ± SE	-0.08 ± 0.028	-0.01 ± 0.00	-0.02 ± 0.00	-0.02 ± 0.00	-0.01 ± 0.00	-0.01 ± 0.00
	t-val	-3.01	-8.68	-8.21	-4.65	-4.32	-2.20
	(*p*-val)	(0.002)**	(<.001)***	(<0.001)***	(<.001)***	(<.001)***	(0.028)*
Fix (s)	β ± SE				-0.00 ± 0.00	0.00 ± 0.00	
	t-val	-	-	-	-0.62	0.58	-
	(*p*-val)				(0.532)	(0.561)	
Stim (s)	β ± SE					-0.00 ± 0.00	
	t-val	-	-	-	-	-0.02	-
	(*p*-val)					(0.985)	
Mask (s)	β ± SE				-0.00 ± 0.00	-0.00 ± 0.00	
	t-val	-	-	-	-1.11	-0.78	-
	(*p*-val)				(0.267)	(0.436)	
Sess.	β ± SE	41.97 ± 19.46	0.71 ± 0.84	0.96 ± 0.70	1.41 ± 0.920	0.86 ± 0.78	-0.87 ± 1.14
	t-val	2.16	0.85	1.37	1.53	1.10	-0.77
	(*p*-val)	(0.031)*	(0.396)	(0.172)	(0.125)	(0.273)	(0.443)

β: estimated regression coefficients for fixed effects. SE: estimated standard error of the regression coefficients.

Val.: valence; **Inf.**: information; **Fix.**: fixation duration; **Stim.**; stimulus display duration; **Sess**.: session number.

~ p <.1; * p <.05; ** p <.01; *** p <.001

B.4 Effect of Experiments

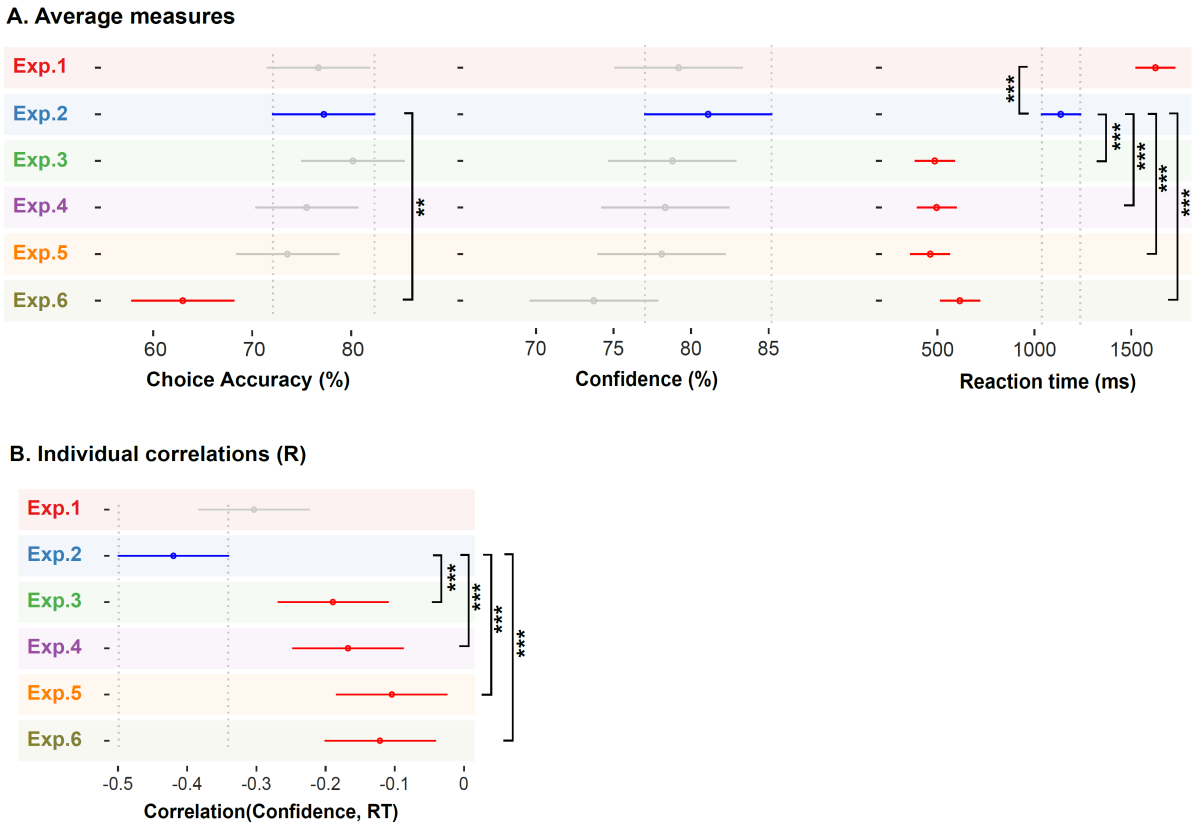


Figure B.4.1. Post-hoc analysis results. The post-hoc tests were performed using the *multcompare* function in Matlab, that implements Tukey's honestly significant difference criterion to deal with multiple comparisons. Means of Choice Accuracy, Confidence and RT for each experiment are represented by a circle. The 95% confidence interval (CI) is represented by a line extending out of the circle. The **post-hoc analysis** compared each experiment with experiment 2 (which served as a basis to design experiments 3-6). **(A)** Left: **Accuracy**. Middle: **Confidence**. Right: **Response times**. **(B) Correlation between confidence and RT**. Experiment 2 is represented by a blue dot ± 95% CI error bar. Experiments represented with a red (resp. grey) dot ± 95% CI error bars were significantly (resp. not significantly) different from Experiment 2.
~ p <.1; * p <.05; ** p <.01; *** p <.001

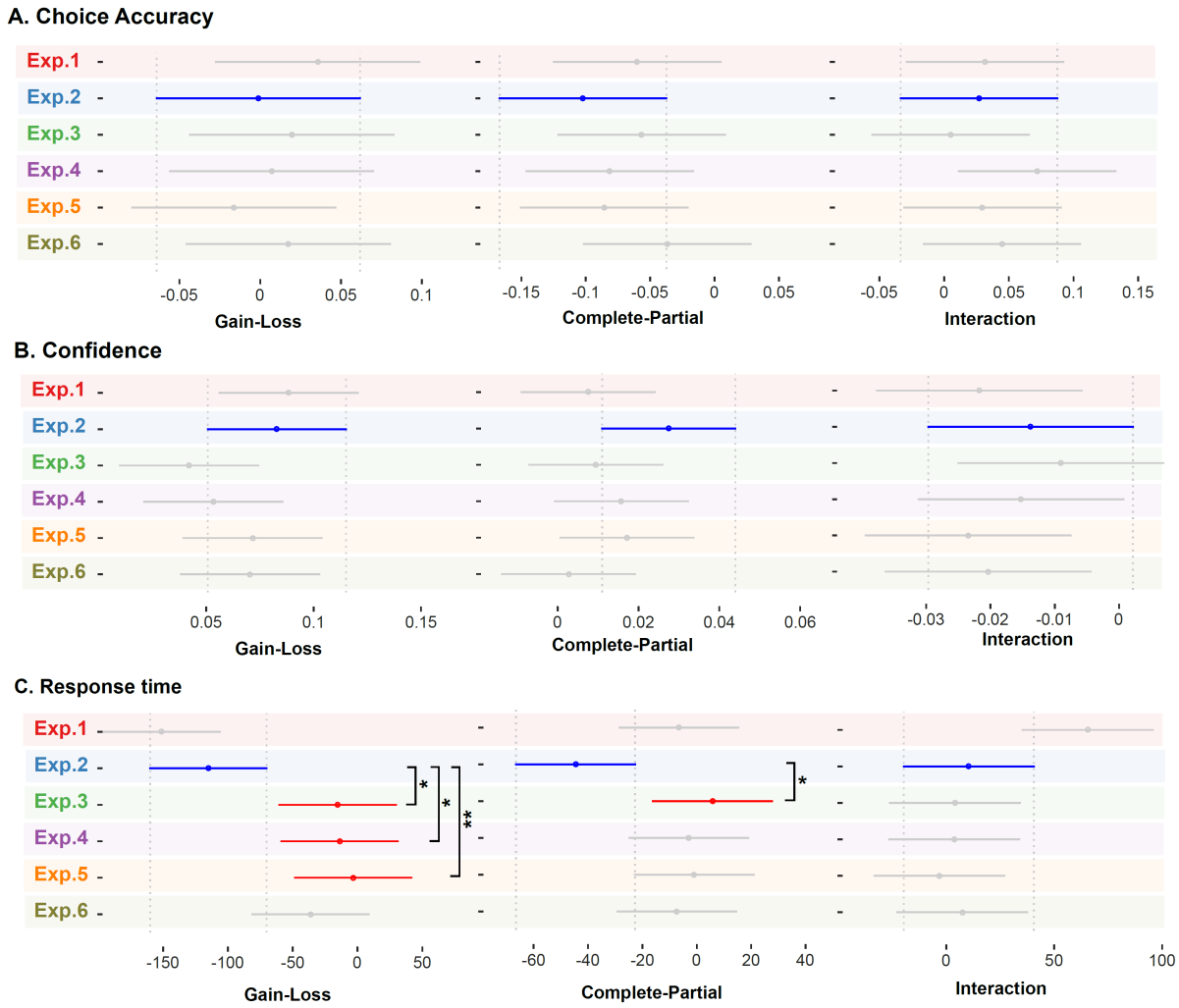


Figure B.4.2. Post-hoc analysis with multiple comparison test. The post-hoc tests were performed with the *multcompare* function in Matlab, and implement Tukey's honestly significant difference criterion to deal with multiple comparisons. The mean effects of the experimental treatments (Valence, Information and their interaction) for each experiment are represented by a circle. The 95% confidence interval (CI) is represented by a line extending out of the circle. The **post-hoc analysis** compared each experiment with experiment 2 (whose experimental design served as a basis for experiments 3-6). **(A)** Effects of Valence (gain-loss), Information (complete – partial) and their Interaction on Accuracy. **(B)** Effects of Valence (gain-loss), Information (complete – partial) and their Interaction on Confidence. **(C)** Effects of Valence (gain-loss), Information (complete – partial) and their Interaction on RT.
Experiment 2 is represented with a blue dot ± 95% CI error bar. Experiments represented with a red (resp. grey) dot ± 95% CI error bars were significantly (resp. not significantly) different from Experiment 2.
~ p<.1; * p <.05; ** p <.01; *** p <.001

B.5 Robust regression analyses

Table B.5. Estimated coefficients from inter-individual robust regressions.

		Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5	Exp. 6	All
Intercept	β ± SE	8.02 ± 2.15	2.94 ± 1.29	2.16 ± 0.95	3.54 ± 1.37	6.76 ± 1.81	3.59 ± 1.95	5.62 ± 0.69
	t-val	3.72	2.27	2.27	2.58	3.73	2.41	8.18
	(*p*-val)	(0.002)**	(0.037)*	(0.038)*	(0.020)*	(0.002)**	(0.028)*	(<.001)***
Slope	β ± SE	-0.00±0.01	-0.03 ± 0.01	-0.02 ± 0.04	-0.06 ± 0.06	0.17 ± 0.18	-0.10 ± 0.03	-0.02 ± 0.01
	t-val	-0.27	-3.55	-0.46	-0.97	0.81	-3.35	-3.75
	(*p*-val)	(0.793)	(0.003)**	(0.662)	(0.35)	(0.368)	(0.004)**	(<.001)***

For each individual, we estimated the net effect of valence on RT and confidence, by computing the averaged difference of these behavioral measures in the gain versus loss contexts. For analyses restricted to a single experiment, we used robust regressions to decrease the vulnerability of our estimates in the relatively small samples (n=18). For the combined analysis (n=108; yellow area in the table), simple and robust regressions gave similar results, and we only report here the results of the simple regression.

β: estimated regression coefficient. SE: estimated standard error of the regression coefficient.

~ p<.1; * p <.05; ** p <.01; *** p <.001

APPENDIX C: Supplementary Material for Chapter 4

C.1 Demographics and behavior results (with model-free analyses)

Table C.1.1. Demographics and behavior.

Gender	**Age**	**Performance (%)**	**Confidence (%)**	**RT (ms)**
M/F	mean ± STD	mean ± SEM	mean ± SEM	mean ± SEM
17/23	22.69±4.44	78.31 ± 1.58	78.71 ± 1.30	715.76 ±11.17

M: Male**; F:** Female**; STD:** standard deviation**; SEM:** standard error of the mean

Table C.1.2. Repeated measures ANOVA results reported for learning performance (i.e., choice accuracy rate), confidence and reaction time (RT).

Performance			**Confidence**			**RT**		
F(1,39)			F(1, 39)			F(1, 39)		
[η^2]			[η^2]			[η^2]		
(*p*-val.)			(*p*-val.)			(*p*-val.)		
val.	inf.	val.×inf.	val.	inf.	val.×inf.	val.	inf.	val.×inf.
0.00	22.05	0.01	36.56	6.76	9.62	4.77	0.31	0.97
[0.00]	[0.07]	[0.00]	[0.10]	[0.01]	[0.00]	[0.01]	[0.00]	[0.00]
(0.967)	(<.001)***	(0.906)	(<.001)***	(0.013)*	(0.004)**	(0.035)*	(0.578)	(0.332)

Val.: outcome valence (gain/loss); **inf.**: information (partial/complete)
~ *p*<.10; * *p* <.05; ** *p* <.01; *** *p* <.001

Table C.1.3. Correlation between confidence and RT.

	Gain partial	Loss partial	Gain Complete	Loss complete	Overall
mean ± SEM	-0.18 ± 0.03	-0.21 ± 0.03	-0.18 ± 0.03	-0.20 ± 0.03	-0.19 ± 0.02
t(39)	-5.56	-6.13	-6.63	-5.86	-8.21
(*p*-val)	(<.001)***	(<.001)***	(<.001)***	(<.001)***	(<.001)***

The correlation between confidence and learning performance (i.e., choice accuracy rate) was performed at the session level using Pearson's R, then averaged at the individual level. Reported statistics correspond to a random-effects analysis (one sample t-test) performed at the population level.
SEM: standard error of the mean. **t**: Student t-value.
~ *p*<.10; * *p* <.05; ** *p* <.01; *** *p* <.001

Table C.1.4. Estimated coefficients from inter-individual robust regressions.

		Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5	Exp. 6	**fMRI**
Intercept	β ± SE	8.02 ± 2.15	2.94 ± 1.29	2.16 ± 0.95	3.54 ± 1.37	6.76 ± 1.81	3.59 ± 1.95	5.02 ± 0.84
	t-val	3.72	2.27	2.27	2.58	3.73	2.41	5.97
	(*p*-val)	(0.002)**	(0.037)*	(0.038)*	(0.020)*	(0.002)**	(0.028)*	(<.001)***
Slope	β ± SE	-0.00±0.01	-0.03 ± 0.01	-0.02 ± 0.04	-0.06 ± 0.06	0.17 ± 0.18	-0.10 ± 0.03	-0.01± 0.01
	t-val	-0.27	-3.55	-0.46	-0.97	0.81	-3.35	-0.97
	(*p*-val)	(0.793)	(0.003)**	(0.662)	(0.35)	(0.368)	(0.004)**	(0.339)

For each individual, we estimated the net effect of valence on RT and confidence, by computing the averaged difference of these behavioral measures in the gain versus loss contexts. The fMRI experiment with larger sample size (n=40; yellow area) revealed the similar results as we found from Chapter 3 (Experiment 1-6, ns= 18).
β: estimated regression coefficient. SE: estimated standard error of the regression coefficient.
~ p<.10; * p <.05; ** p <.01; *** p <.001

Table C.1.5. Parameters and correlation analyses.

A. Mean and stander deviation of parameters for each model within model space.

	β	α_c	α^+	α^-	α_u	α_v	w
RW	2.25±1.14	0.51±0.25	-	-	-	-	-
ABS	2.73±1.37	0.42±0.27	-	-	0.25±0.25	-	-
REL	5.69±2.52	0.23±0.19	-	-	0.17±0.10	0.32±0.34	-
REL_α	3.98±1.46	-	0.45±0.25	0.13±0.12	-	0.42±0.43	-
REL_w	7.48±3.95	0.19±0.17	-	-	0.13±0.10	0.39±0.38	0.64±0.33
$REL_{\alpha,w}$	4.78±2.15	-	0.38±0.24	0.12±0.10	-	0.33±0.37	0.64±0.33

B. Correlations between parameters

	β	α^+	α^-	α_v	w
α^+	-0.63 (p <.001)***	-	-	-	-
α^-	0.049 (p = .764)	0.37 (p =.020)*	-	-	-
α_v	-0.05 (p = .758)	0.03 (p =.856)	-0.12 (p =.471)	-	-
w	0.38 (p = .017)**	-0.32 (p =.045)*	-0.11 (p =.500)	-0.05 (p =.759)	-
Confidence bias	**-0.33 (p =.041)***	**0.09 (p =.578)**	**-0.31 (p =.0501)~**	**-0.27 (p =.091)~**	**-0.25 (p =.123)**

The correlation between parameters from the winning model (RELα,w) and average confidence bias. It is worth to note that even though the parameters were highly correlated, the winning model was reliable enough to pass the parameter and model recovery tests.

~ p<.10; * p <.05; ** p <.01; *** p <.001

A. Model identifiability **B. Parameter recovery ($REL_{\alpha,w}$)**

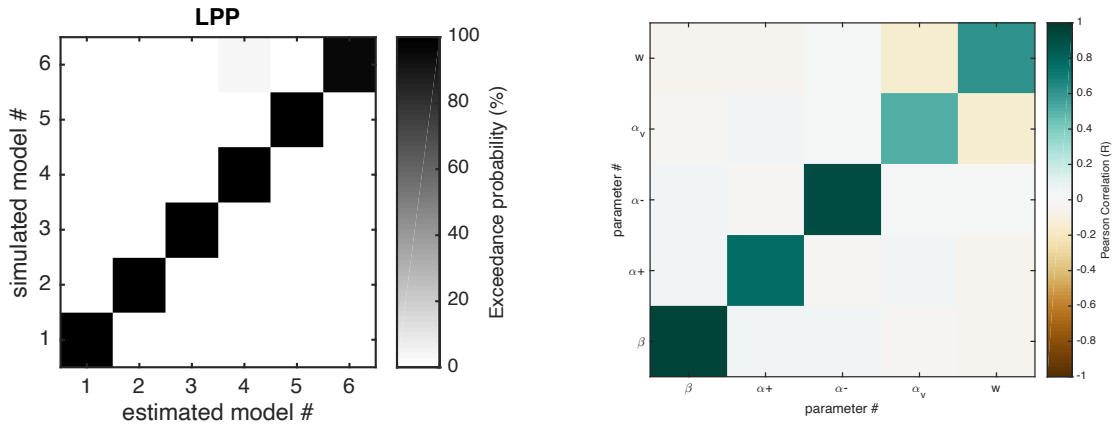


Figure C.1.1. Model and parameter recovery. (A) Average exceedance probability confusion matrix for model identifiability analysis. We simulated data from 40 synthetic participants with each of six models within model space. We then identified the most probable model that generating simulated data using Bayesian model selection. Afterward, the estimated model was compared to simulated model to illustrate the accuracy of attribution. This procedure was repeated five times. **(B)** Parameter recovery analysis. The confusion matrices represent summary statistics of the correlations between parameters, estimated over 40-subject simulations, and averaged over the five simulations. Diagonal: correlations between simulated and estimated parameters. Off diagonal: cross correlation between estimated parameters.

C.2 fMRI analysis and results

Table C.2.1. Summary of GLM.

Events	Symbols	Choice	Confidence Rating	Outcome
regressors	GP_onset PM: confidence LP_onset PM: confidence GC_onset PM: confidence LC_onset PM: confidence	Onset PM: RT	Onset PM: Distance PM: Rating's RT	GP_onset PM: good/bad (1/0) LP_onset PM: good/bad (1/0) GC_onset PM: good/bad (1/0) LC_onset PM: good/bad (1/0)

The GLM consisted of 10 event onsets and 12 corresponded parametric modulators. Specifically, the symbol presentation, choice, confidence rating and outcome onsets were used. Symbol presentation and outcome onsets were selectively divided by four conditions, so that we were able to investigate context-dependent confidence and outcome encoding. The good outcome referred to 1€ and -0.1€ in gain and loss contexts, respectively. The bad outcome referred to 0.1€ and -1€ in gain and loss contexts, respectively. Distance denoted the differences between starting point and stating point.

GP: Gain partial; **LP**: Loss partial; **GC**: Gain Complete; **LC**: Loss Complete

PM: parametric modulator.

Table C.2.2. Regions encoded outcome and confidence during outcome and symbol presentation, respectively.

A. Outcome

Contrast	Label	x	y	z	# of activation	t-value
Outcome (Positive)	VS	9	14	10	692	8.58
		-9	50	-7		6.74
		-12	14	-10		6.42
	PCC	18	-48	62	238	5.69
		15	-34	68		4.60
		9	-55	65		4.57
	Cingulum_Mid	15	-28	44	74	5.45
		15	-16	47		4.01
	Angular gyrus	-42	-73	32	46	5.38
		-42	-64	29		4.48
	STG (superior temporal gyrus)	66	-31	17	252	5.32
		66	-25	11		5.27
		66	-7	11		4.69
	cerebelum	27	-58	-43	161	5.28
		45	-67	-40		4.67
		42	-49	-43		4.46
	Cuneus	18	-88	26	159	5.01
		15	-94	11		4.89
		-6	-88	26		3.99
	Putamen	30	-13	-1	151	4.89
		30	-7	8		4.75
		30	-4	-1		4.67
	Middle occipital gyrus	33	-70	5	143	4.85
		15	-85	-13		4.64
	Caudate Nucleus	21	5	23	66	4.83
	Hippocampus	30	-16	-19	62	4.81
		24	-7	-25		3.77
	Primary motor	-18	-28	62	84	4.50
		-15	-34	68		4.14
		-21	-28	53		4.01
Contrast	**Label**	**x**	**y**	**z**	**# of activation**	**t-value**
Outcome (Negative)	Insula	39	20	-10	199	6.89
		36	23	5		5.79
		30	20	-16		4.26
	dmPFC	6	35	47	269	6.52
		6	44	35		6.43
		-3	17	56		3.84
	Right IFG	48	17	29	275	6.30
		48	26	23		5.70
		39	8	26		4.78
	Left IFG	-45	5	29	69	4.19
		-51	29	29		4.08
		-39	17	23		3.75

B. Confidence

Contrast	Label	x	y	z	# of activation	t-value
Confidence (Positive)	MTL	-60	-25	-4	72	6.38
		-51	-34	5		4.48
	vmPFC	0	41	2	258	5.97
		-9	50	-4		5.49
		-6	65	5		5.38
	Angular gyrus BA18	-12	-82	-10	204	5.67
		-6	-88	5		3.96
		-3	-94	-4		3.71
	Angular gyrus BA39	-48	-64	26	89	4.12
		-57	-64	14		4.00
		-57	-55	17		3.87
	Precentral gyrus	-27	-25	59	43	4.09
		-36	-16	56		3.90
		-30	-13	65		3.56
Contrast	**Label**	**x**	**y**	**z**	**# of activation**	**t-value**
Confidence (Negative)	IFG	48	8	23	274	6.85
		45	32	11		5.77
		48	26	23		5.15
	Caudate	-9	11	2	59	4.39
		-15	20	2		4.02
	Right Insula	33	23	-4	138	5.67
		15	14	2		4.52
		6	8	5		4.10
	Left Insula	-30	23	-4	53	4.23
	dmPFC	-3	26	41	86	5.44
		9	23	41		4.76
	Right FFA	27	-40	-16	282	5.98
		45	-43	-19		5.91
		48	-55	-13		4.79
	Left FFA	-42	-46	19	91	5.15
		-42	-61	-10		4.09
		-42	-70	-7		4.03
Contrast	**Label**	**x**	**y**	**z**	**# of activation**	**t-value**
Confidence (Gain>Loss)	N.S.	-	-	-	-	-
Confidence (Complete>Partial)	N.S.	-	-	-	-	-

k= p<.001 whole-brain cluster-defining height threshold at uncorrected

C.3 Instruction

Dear participant,

First, thank you for participating in our experiment. Before the experiment starts, it is important that you are informed about the procedures. Therefore, we ask you to read this information letter carefully. Please do not hesitate to ask for clarification about this text or the general procedure, if anything is unclear. Your experimenter is happy to answer your questions.

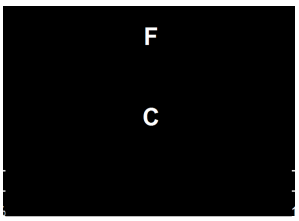
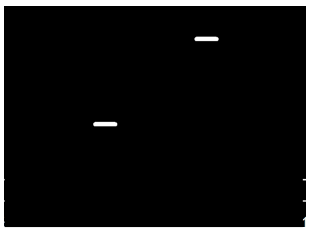
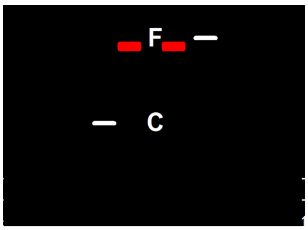
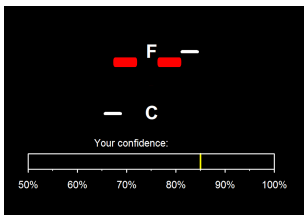
Goal of the study
The purpose of the experiment is to investigate how people learn the probabilities of winning and losing money. The experiment will be conducted inside an MRI scanner. We will track your choices and brain responses throughout the experiment.

Procedure
During this game, you are asked to make repeated choices between two symbols shown to you on the computer screen. All symbols carry a certain value and the two symbols that are displayed simultaneously are not equivalent. One is **on average** more advantageous than the other, either because it brings big gains more often, or because it brings fewer big losses when compared to the other symbol of the pair. The result of your choice may be that:

- you earn money (+1 € or + 0.10 €)
- you lose money (-1€ or - 0.10 €)

The goal of the game is to win as much money as possible, even if avoiding losses is not possible at all times. Please note that gains and losses never occur together in one trial, meaning that symbol pairs can lead to only wins, or only losses. With losing symbol pairs, the goal is to get the small loss more frequently than the big loss; with winning symbol pairs, the goal is to get the big gain more often.

The table on the following page informs you of the type of decisions you will be making in the experiment. As you can see, your decisions proceed in six stages → see next page.

Stage 1.	You will be shown two symbols, one on the top and one on the bottom. In this stage, we ask you to evaluate these symbols and decide which one you prefer, but without pressing a button for now.	
Stage 2.	The symbols will then disappear for a short while. At this point you do not know which button corresponds to which symbol, as this will be reassigned by a random mechanism for each trial. Due to this, you need to remember the location of the symbol that you selected until the next stage.	
Stage 3.	Two white bars now appear to inform you which button to press to select your preferred symbol. These occur in a location next to the symbols, but the symbols are not shown anymore. Although you cannot see the symbols when you press the button, the location of the symbols is unchanged in a given trail. As shown in the figure, if the white bar next to the location of your chosen symbol is on the right, the right button selects this symbol. If the white bar is on the left, pressing the left button selects the symbol.	
Stage 4.	After selection of your preferred symbol you will receive confirmation via two red bars (-) that appear in the location of the symbol you selected.	
Stage 5.	In this stage, you are asked to indicate how confident you are on a scale from 50% to 100%. 50% means that you made your choice completely at random and you do not know which of the symbols has the best value; 100% means that you are absolutely certain you chose the symbol that has the best average value of the two. Your final confidence rating should represent your estimated probability that you correctly chose the symbol that is the relative advantageous of the pair, because your confidence accuracy can lead to additional bonus earnings for you. You can move the bar by pressing the left and right buttons. To confirm your decision press the extreme right button.	

Stage 6.	You will receive feedback from chosen option. We ask you to try to optimize your winnings based on this feedback. Sometimes to help you, the result corresponding to the choice of the other symbol (not chosen) will also be provided, sometimes it won't.	

Additional information:

- We ask you to try and make a choice on every trial as soon as possible.

- Because the rewards are probabilistic, the more **advantageous symbol** will, in some trials, give a smaller gain (or a bigger loss) than the other symbol. However, the more advantageous symbol will clearly provide higher gains and lower losses on average.

- The symbol pairs are fixed in a run, meaning that you will always see the same symbols paired with each other. The location in which a symbol appears (top or bottom) does not change the value of the symbol.

- In the experiment, outcomes associated with the symbol of your choice **are not influenced by your confidence rating**. However, you can win a confidence bonus, which is highest if your confidence matches your actual probability of having chosen the most advantageous symbol (this will be explained in details after practice). It is therefore important to be **as truthful and as accurate as possible when estimating your confidence** during the learning task.

- Before the task starts you will go through a few practice trials to familiarize yourself with the task and its pace. Please make sure to ask your experimenter any questions during or after the practice trials. After the practice trials, an example of how the confidence bonus works will be shown. When everything is clear to you, you can start with the main experiment, which consists of 3 parts, each should take about 15-20 minutes (80 trials for each part). Each part contains different symbols and you will re-learn all values, but your task is the same. After these 3 parts, you will do an additional task (112 trials), for which you are given separate instructions before the start. All tasks combined will take about one hour and 15 minutes.

- At the end of the session, the experimenter will tell you the total of your winnings, which is based on all the choices that you made throughout the experiment.

Potential discomfort and risks

MRI is a safe technique that is already being used in hospitals and research centers for over 20 years. A strong magnetic field is active in the space in which the MRI-research takes place. It is therefore important that you comply with the safety instructions of the operating staff. Prior to the experiment, you will be asked to fill in a "MRI screening" form. This will check whether

you can safely participant in MRI research; therefore you must complete this form completely and truthfully. If you follow the safety instructions of the operating staff and complete the MRI screening truthfully, there are no known health risks associated with MRI. The signed form will be archived for a maximum of two years.

In the supplement *Information MRI research* you can find additional information about the risk factors and possible discomfort associated with MRI research. Read this supplement carefully. In summary, for pure (non-medical) research the following characteristics apply that you may experience as unpleasant:

Narrow space. In the MRI-scanner your head and part of your body will lie in a relatively narrow tube during the experiment (70 minutes). This does not cause any problems for most people, but if you are slightly claustrophobic, we ask you to report this to the researcher. If you suffer from severe claustrophobia it is better not to participate in the experiment

Limited freedom of movement. During the scanning you should not move your head, and that can cause a feeling of cramping or stiffening. In some case this causes a headache. Note that this is not due to the magnetic field itself.

Loud sound. During scanning, the device makes a loud, sometimes thumping, sound; this is normal. Since you will wear adequate hearing protection (this will be provided by the research center), this sound is not harmful to the hearing. Therefore, make sure that the hearing protection that you receive during the experiment is properly used.

It is important to know that you are always keep in touch with the researcher during a scan. By pressing a button you can communicate with the researcher as all times. The researcher will then talk to you via the intercom and, if requested, will immediately remove you from the scanner.

Insurance
As with any research at the University of Amsterdam, a standard liability insurance applies.

Your privacy is guaranteed
Your personal information (for instance name, date of birth, address) remains confidential and will never be shared with third parties. Research data that are published in scientific journals will be anonymous and cannot be traced back to you as an individual.

Sharing anonymized research data
I agree that my completely anonymized data can be made publicly accessible.

Explanation of making the anonymized research data publicly accessible
All research data collected in the current study can be used in other, future research. Such future research can focus on questions that are not related to the current study.

The research data will be shared anonymously, but freely accessible on the internet via a public database. The research data that will be shared will not contain any personal information such as name, address, date of birth, date of participation and facial features or other information that would make it possible for others to directly identify you. The anonymity described above (by means of separating the research data from the personal information) is no longer guaranteed when a third party (e.g. the participant themselves or an institution that has your MRI data) shares your MRI data including the personal information.

If you do not agree to make your anonymized research data publicly accessible, you cannot participate in this study.

Incidental findings Spinoza Centre Roeterseiland

There is a small chance that we find an abnormality in your brain during an MRI experiment. Often, these abnormalities are small deviations or normal variances; but in certain case this could be severe (such as a brain tumour). If this is the case, the information will be examined by a radiologist and sent it to your general practitioner. Your general practitioner will contact you in that case. In case the researcher sends the scans to the radiologist of the MRI Centrum Amsterdam, your scans have to be stored for at least 15 years at the MRI Centrum Amsterdam. This is done in accordance with the law.

Hereby we want to remind you that the technician / researcher is not medically adept nor trained to detect all forms of brain abnormalities. Furthermore, it is technically not possible to detect all forms of brain damages on the scans acquired. In order to participate in an MRI experiment, it is mandatory to agree with this procedure beforehand by providing the name, telephone number and address of your general practitioner and your social security number (BSN).

Access to own research data

It is not possible to view your own research data (MRI or otherwise) after participating in a research project at the Spinoza Centre.

Compensation

Throughout the experiment we will use laboratory currency (MU) that will be converted to Euros after all procedures are completed (exchange rate: 1 MU = 0.3 Euros). You will receive a show-up fee of 20 Euros, to which additional winnings and losses based on your decisions will be applied. There are two sources for the additional income:

1. Your learning performance, which in past experiments has led to average additional payments of ca. 7 Euros;
2. Confidence bonus.

The computer will keep track of your wins and losses during the game. A confidence bonus can be won (as mentioned before). The mechanism used to determine the bonus is visualized in an example after your practice trials. After the last task, the same mechanism will tell you which bonuses you earned and how much your final payout is. It will use random trials from your sessions, so it is in your best interest to be mindful of your answers in all trials.

Voluntary participation

There are no consequences if you decide now not to participate in this study. During the experiment, you are free to stop participating at any moment without giving a reason for doing so. You can request that your research data be deleted within 7 days of the investigation.

Further information

Should you have questions about this study at any given moment, please contact the responsible researchers:

Dr. Jan Engelmann
Tel: (0) 205 255 5651
Email: j.b.engelmann@uva.nl
Roeterstraat 11,
1018 WB, Amsterdam.

Chih-Chung Ting
Email: c.ting@uva.nl
UvA, REC-E7.23,
Roeterstraat 11,
1018 WB, Amsterdam.

Formal complaints about this study can be addressed to the Faculty (FMG) Ethics Review Board of the University of Amsterdam:

Dr. Wery van den Wildenberg
Tel: 020-5256686.
Email: w.p.m.vandenwildenberg@uva.nl
Fmg-UvA, REC-G1.10,
Nieuwe Achtergracht 129 B,
1018 WS Amsterdam

CONSENT FORM

This form belongs to the information letter that you received and that describes the research in which you participate. By signing this form, you declare that you understand the nature and methods of this study as described in the information letter. Furthermore, by signing you agree with the experimental procedure as described in the information letter.

If you have questions about this study, or you wish to receive further information regarding the research at any given moment, please contact the responsible researchers:

Dr. Jan Engelmann
Tel: (0) 205 255 5651
Email: j.b.engelmann@uva.nl
Roeterstraat 11,
1018 WB, Amsterdam.

Chih-Chung Ting
Email: c.ting@uva.nl
UvA, REC-E7.23,
Roeterstraat 11,
1018 WB, Amsterdam.

Formal complaints about this study can be addressed to the Ethics Review Board:

Dr. Wery van den Wildenberg
Tel: 020-5256686.
Email: w.p.m.vandenwildenberg@uva.nl
Fmg-UvA, REC-G1.10,
Nieuwe Achtergracht 129 B,
1018 WS Amsterdam

Signed in duplicate

[PARTICIPANT]

- *I am 16 or older.*
- *I have read and understood the information letter.*
- *I agree to participate in this study and I agree with the use of the data that are collected.*
- *I reserve the right to withdraw my consent at any moment without providing any reason.*
- *I reserve the right to withdraw my participation from the study at any moment without providing any reason.*
- I **do/do not** give permission to keep my contact details so that we can send you information about possible follow-up research.
  *Strike out what is not applicable. N.B. This of course does NOT mean that you already agree to participate in the research about which you received information. You will be able to decide about this in due time without this having any consequences for you.

.....

....

participant name *participant signature* *date*

[**RESEARCHER**]

- *I informed the participant about the research;*
- *I am willing to answer any possible questions about the research to the best of my ability.*

.....

....

researcher name *researcher signature* *date*

Bibliography

Abe, N., & Greene, J. D. (2014). Response to Anticipated Reward in the Nucleus Accumbens Predicts Behavior in an Independent Test of Honesty. *Journal of Neuroscience*, *34*(32), 10564–10572.

Abraham, A., & Hermann, C. (2015). Biases in probabilistic category learning in relation to social anxiety. *Frontiers in Psychology*, *6*.

Bailey, J. E., Dawson, G. R., Dourish, C. T., & Nutt, D. J. (2011). Validating the nhalation of 7.5% CO2 in healthy volunteers as a human experimental medicine: A model of generalized anxiety disorder (GAD). *Journal of Psychopharmacology*, *25*(9), 1192–1198.

Bandelow, B., & Michaelis, S. (2015). Epidemiology of Anxiety Disorders in the 21st Century. *Clinical Research*, *17*(3), 9.

Bang, D., & Fleming, S. M. (2018). Distinct encoding of decision confidence in human medial prefrontal cortex. *Proceedings of the National Academy of Sciences*, *115*(23), 6082–6087.

Barrett, L. F., Bliss-Moreau, E., Duncan, S. L., Rauch, S. L., & Wright, C. I. (2007). The amygdala and the experience of affect. *Social Cognitive and Affective Neuroscience*, *2*(2), 73–83.

Barto, A. G., & Sutton, R. S. (1997). Reinforcement Learning in Artificial Intelligence. In *Advances in Psychology* (Vol. 121, pp. 358–386). Elsevier.

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*, *76*, 412–427.

Bavard, S., Lebreton, M., Khamassi, M., Coricelli, G., & Palminteri, S. (2018). Reference-point centering and range-adaptation enhance human reinforcement learning at the cost of irrational preferences. *Nature Communications*, *9*(1), 4503.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, *56*(6), 893–897.

Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, *8*(1), 77–100.

Becker, G. M., Degroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, *9*(3), 226–232.

Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring Utility by a Single-Response Sequential Method. *Behavioral Science*, *9*(3), 226–232.

Behar, E., DiMarco, I. D., Hekler, E. B., Mohlman, J., & Staples, A. M. (2009). Current theoretical models of generalized anxiety disorder (GAD): Conceptual review and treatment implications. *Journal of Anxiety Disorders*, *23*(8), 1011–1023.

Berghorst, L. H., Bogdan, R., Frank, M. J., & Pizzagalli, D. A. (2013). Acute stress selectively reduces reward sensitivity. *Frontiers in Human Neuroscience*, *7*.

Berner, E. S., & Graber, M. L. (2008). Overconfidence as a Cause of Diagnostic Error in Medicine. *The American Journal of Medicine*, *121*(5), S2–S23.

Bishop, S. J., & Gagne, C. (2018). Anxiety, Depression, and Decision Making: A Computational Perspective. *Annual Review of Neuroscience*, *41*(1), 371–388.

Boldt, A., Blundell, C., & De Martino, B. (2017). *Confidence modulates exploration and exploitation in value-based learning* [Preprint]. Neuroscience.

Bolton, S., & Robinson, O. J. (2017). The impact of threat of shock-induced anxiety on memory encoding and retrieval. *Learning & Memory*, *24*(10), 532–542.

Bor, D., Schwartzman, D. J., Barrett, A. B., & Seth, A. K. (2017). Theta-burst transcranial magnetic stimulation to the prefrontal or parietal cortex does not impair metacognitive visual awareness. *PLOS ONE*, *12*(2), e0171793.

Boureau, Y.-L., & Dayan, P. (2011). Opponency Revisited: Competition and Cooperation Between Dopamine and Serotonin. *Neuropsychopharmacology*, *36*(1), 74–97.

Bradley, B. P., Mogg, K., & Millar, N. H. (2000). Covert and overt orienting of attention to emotional faces in anxiety. *Cognition and Emotion*, *14*(6), 789–808.

Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2002). Region of Interest Analysis Using an SPM Toolbox. *NeuroImage*, *16*(2), 769–1198.

Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X., & Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature Neuroscience*, *18*(4), 590–596.

Büchel, C., Brassen, S., Yacubian, J., Kalisch, R., & Sommer, T. (2011). Ventral striatal signal changes represent missed opportunities and predict future choice. *NeuroImage*, *57*(3), 1124–1130.

Buckert, M., Schwieren, C., Kudielka, B. M., & Fiebach, C. J. (2014). Acute stress affects risk taking but not ambiguity aversion. *Frontiers in Neuroscience*, *8*.

Buhr, K., & Dugas, M. J. (2006). Investigating the construct validity of intolerance of uncertainty and its unique relationship with worry. *Journal of Anxiety Disorders*, *20*(2), 222–236.

Camerer, C., & Lovallo, D. (1999). Overconfidence and Excess Entry: An Experimental Approach. *American Economic Review*, *89*(1), 306–318.

Carleton, R. N., Mulvogue, M. K., Thibodeau, M. A., McCabe, R. E., Antony, M. M., & Asmundson, G. J. G. (2012). Increasingly certain about uncertainty: Intolerance of uncertainty across anxiety and depression. *Journal of Anxiety Disorders*, *26*(3), 468–479.

Carlson, J. M., Greenberg, T., Rubin, D., & Mujica-Parodi, L. R. (2011). Feeling anxious: Anticipatory amygdalo-insular response predicts the feeling of anxious anticipation. *Social Cognitive and Affective Neuroscience*, *6*(1), 74–81.

Cartoni, E., Balleine, B., & Baldassarre, G. (2016). Appetitive Pavlovian-instrumental Transfer: A review. *Neuroscience & Biobehavioral Reviews*, *71*, 829–848.

Cavanagh, J. F., Bismark, A. W., Frank, M. J., & Allen, J. J. B. (2019). Multiple dissociations between comorbid depression and anxiety on reward and punishment processing: Evidence from computationally informed EEG. *Computational Psychiatry*, *3*, 1–17.

Cavanagh, J. F., Frank, M. J., & Allen, J. J. B. (2011). Social stress reactivity alters reward and punishment learning. *Social Cognitive and Affective Neuroscience*, *6*(3), 311–320.

Cerliani, L., Thomas, R. M., Jbabdi, S., Siero, J. C. W., Nanetti, L., Crippa, A., Gazzola, V., D'Arceuil, H., & Keysers, C. (2012). Probabilistic tractography recovers a rostrocaudal trajectory of connectivity variability in the human insular cortex. *Human Brain Mapping*, *33*(9), 2005–2034.

Charness, G., Garcia, T., Offerman, T., & Villeval, M. C. (2020). Do measures of risk attitude in the laboratory predict behavior under risk in and outside of the laboratory? *Journal of Risk and Uncertainty*.

Charpentier, A., Elie, R., & Remlinger, C. (2020). Reinforcement Learning in Economics and Finance. *arXiv preprint. arXiv:2003.10014*.

Charpentier, C. J., Aylward, J., Roiser, J. P., & Robinson, O. J. (2017). Enhanced risk aversion, but not loss aversion, in unmedicated pathological anxiety. *Biological Psychiatry*, *81*(12), 1014–1022.

Charpentier, C. J., Martino, B. D., Sim, A. L., Sharot, T., & Roiser, J. P. (2016). Emotion-induced loss aversion and striatal-amygdala coupling in low-anxious individuals. *Social Cognitive and Affective Neuroscience*, *11*(4), 569–579.

Chen, C., Takahashi, T., Nakagawa, S., Inoue, T., & Kusumi, I. (2015). Reinforcement learning in depression: A review of computational research. *Neuroscience & Biobehavioral Reviews*, *55*, 247–267.

Cheng, C., & Cheung, M. W. L. (2005). Psychological Responses to Outbreak of Severe Acute Respiratory Syndrome: A Prospective, Multiple Time-Point Study. *Journal of Personality*, *73*(1), 261–285.

Chua, E. F., Pergolizzi, D., & Weintraub, R. R. (2014). The Cognitive Neuroscience of Metamemory Monitoring: Understanding Metamemory Processes, Subjective Levels Expressed, and Metacognitive Accuracy. In S. M. Fleming & C. D. Frith (Eds.), *The Cognitive Neuroscience of Metacognition* (pp. 267–291). Springer Berlin Heidelberg.

Chua, E. F., Schacter, D. L., & Sperling, R. A. (2009). Neural Correlates of Metamemory: A Comparison of Feeling-of-Knowing and Retrospective Confidence Judgments. *Journal of Cognitive Neuroscience*, *21*(9), 1751–1765.

Clark, L., Li, R., Wright, C. M., Rome, F., Fairchild, G., Dunn, B. D., & Aitken, M. R. F. (2012). Risk-avoidant decision making increased by threat of electric shock: Risk avoidance under threat of shock. *Psychophysiology*, *49*(10), 1436–1443.

Clore, G. L., & Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences*, *11*(9), 393–399.

Cohn, A., Engelmann, J., Fehr, E., & Maréchal, M. A. (2015). Evidence for Countercyclical Risk Aversion: An Experiment with Financial Professionals. *American Economic Review*, *105*(2), 860–885.

Colwill, R. M., & Rescorla, R. A. (1988). Associations between the discriminative stimulus and the reinforcer in instrumental learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *14*(2), 155.

Das, E., Kerkhof, P., & Kuiper, J. (2008). Improving the Effectiveness of Fundraising Messages: The Impact of Charity Goal Attainment, Message Framing, and Evidence on Persuasion. *Journal of Applied Communication Research*, *36*(2), 161–175.

Daunizeau, J., Adam, V., & Rigoux, L. (2014). VBA: A Probabilistic Treatment of Nonlinear Models for Neurobiological and Behavioural Data. *PLoS Computational Biology*, *10*(1), e1003441.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704–1711.

De Martino, B., Camerer, C. F., & Adolphs, R. (2010). Amygdala damage eliminates monetary loss aversion. *Proceedings of the National Academy of Sciences*, *107*(8), 3788–3792.

De Martino, B. D., Bobadilla-Suarez, S., Nouguchi, T., Sharot, T., & Love, B. C. (2017). Social Information is Integrated into Value and Confidence Judgments According to its Reliability. *Journal of Neuroscience*, 3880–16.

De Martino, B., Fleming, S. M., Garrett, N., & Dolan, R. J. (2013). Confidence in value-based choice. *Nature Neuroscience*, *16*(1), 105–110.

Desender, K., Opstal, F. V., & Bussche, E. V. den. (2017). Subjective experience of difficulty depends on multiple cues. *Scientific Reports*, *7*, 44222.

Devetag, G., & Warglien, M. (2003). Games and phone numbers: Do short-term memory bounds affect strategic behavior? *Journal of Economic Psychology*, *24*(2), 189–202.

DeVido, J., Jones, M., Geraci, M., Hollon, N., Blair, R. J. R., Pine, D. S., & Blair, K. (2009). Stimulus-reinforcement-based decision making and anxiety: Impairment in generalized anxiety disorder (GAD) but not in generalized social phobia (GSP). *Psychological Medicine*, *39*(07), 1153.

Donoso, M., Collins, A. G. E., & Koechlin, E. (2014). Foundations of human reasoning in the prefrontal cortex. *Science*, *344*(6191), 1481–1486.

Dotan, D., Meyniel, F., & Dehaene, S. (2018). On-line confidence monitoring during decision making. *Cognition*, *171*, 112–121.

Ducharme, W. M., & Donnell, M. L. (1973). Intrasubject comparison of four response modes for "subjective probability" assessment. *Organizational Behavior and Human Performance*, *10*(1), 108–117.

Dugas, M. J., Letarte, H., Rhéaume, J., Freeston, M. H., & Ladouceur, R. (1995). Worry and problem solving: Evidence of a specific relationship. *Cognitive Therapy and Research*, *19*(1), 109–120.

Elliott, R., Agnew, Z., & Deakin, J. F. W. (2008). Medial orbitofrontal cortex codes relative rather than absolute value of financial rewards in humans: Medial OFC and relative value. *European Journal of Neuroscience*, *27*(9), 2213–2218.

Engelmann, J. B., Berns, G. S., & Dunlop, B. W. (2017). Hyper-responsivity to losses in the anterior insula during economic choice scales with depression severity. *Psychological Medicine*, *47*(16), 2879.

Engelmann, J. B., Meyer, F., Fehr, E., & Ruff, C. C. (2015). Anticipatory Anxiety Disrupts Neural Valuation during Risky Choice. *Journal of Neuroscience*, *35*(7), 3085–3099.

Engelmann, J. B., & Hare, T. (2018). Emotions can bias decision-making processes by promoting specific behavioral tendencies. In *The Nature of Emotion: Fundamental Questions* (2nd ed.). Oxford University Press.

Engelmann, J. B., Meyer, F., Fehr, E., & Ruff, C. C. (2015). Anticipatory anxiety disrupts neural valuation during risky choice. *Journal of Neuroscience*, *35*(7), 3085–3099.

Engelmann, J. B., Meyer, F., Ruff, C. C., & Fehr, E. (2019). The neural circuitry of affect-induced distortions of trust. *Science Advances*, *5*(3), eaau3413.

Engelmann, J. B., & Hein, G. (2013). Contextual and social influences on valuation and choice. In *Progress in Brain Research* (Vol. 202, pp. 215–237). Elsevier.

Engelmann, J. B., Maciuba, B., Vaughan, C., Paulus, M. P., & Dunlop, B. W. (2013). Posttraumatic Stress Disorder Increases Sensitivity to Long Term Losses among Patients with Major Depressive Disorder. *PLoS ONE*, *8*(10), e78292.

Engelmann, J. B., Meyer, F., Fehr, E., & Ruff, C. C. (2015). Anticipatory Anxiety Disrupts Neural Valuation during Risky Choice. *Journal of Neuroscience*, *35*(7), 3085–3099.

Erev, Id., & Roth, A. E. (2014). Maximization, learning, and economic behavior. *Proceedings of the National Academy of Sciences*, *111*, 10818–10825.

Erev, Id., & Roth, A. E. (1998). Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria. *American Economic Association*, 35.

Ernst, M., Plate, R. C., Carlisi, C. O., Gorodetsky, E., Goldman, D., & Pine, D. S. (2014). Loss aversion and 5HTT gene variants in adolescent anxiety. *Developmental Cognitive Neuroscience*, *8*, 77–85.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion*, *7*(2), 336–353.

Fleming, S. M., & Daw, N. D. (2017). Self-evaluation of decision-making: A general Bayesian framework for metacognitive computation. *Psychological Review*, *124*(1), 91–114.

Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Phil. Trans. R. Soc. B*, *367*(1594), 1338–1349.

Fleming, S. M., Huijgen, J., & Dolan, R. J. (2012). Prefrontal Contributions to Metacognition in Perceptual Decision Making. *The Journal of Neuroscience*, *32*(18), 6117–6125.

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*.

Fleming, S. M., Maniscalco, B., Ko, Y., Amendi, N., Ro, T., & Lau, H. (2015). Action-specific disruption of perceptual confidence. *Psychological Science*, *26*(1), 89–98.

Fleming, S. M., Putten, E. J. van der, & Daw, N. D. (2018). Neural mediators of changes of mind about perceptual decisions. *Nature Neuroscience*, *21*(4), 617.

Fleming, S. M., Ryu, J., Golfinos, J. G., & Blackmon, K. E. (2014). Domain-specific impairment in metacognitive accuracy following anterior prefrontal lesions. *Brain*, *137*(10), 2811–2822.

Folke, T., Jacobsen, C., Fleming, S. M., & Martino, B. D. (2016). Explicit representation of confidence informs future value-based decisions. *Nature Human Behaviour*, *1*, 0002.

Fontanesi, L., Palminteri, S., & Lebreton, M. (2019). Decomposing the effects of context valence and feedback information on speed and accuracy during reinforcement learning: A meta-analytical approach using diffusion decision modeling. *Cognitive, Affective, & Behavioral Neuroscience*, *19*(3), 490–502.

Frank, M. J., Seeberger, L. C., & Randall C., O. (2004). *By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsonism. 306*, 1940–1943.

Galván, A., & Peris, T. S. (2014). Neural Correlates of Risky Decision Making in Anxious Youth and Healthy Controls: Research Article: fmri of Risky Decisions in Anxious Youth. *Depression and Anxiety*, *31*(7), 591–598.

Garrett, N., & Daw, N. D. (2020). Biased belief updating and suboptimal choice in foraging decisions. *Nature Communications*, *11*(1), 3417.

Geller, E. S., & Whitman, C. P. (1973). Confidence ill stimulus predictions and choice reaction time. *Memory & Cognition*, *1*(3), 361–368.

Giardini, F., Coricelli, G., Joffily, M., & Sirigu, A. (2008). Overconfidence in Predictions as an Effect of Desirability Bias. In P. M. Abdellaoui & P. D. J. D. Hey (Eds.), *Advances in Decision Making Under Risk and Uncertainty* (pp. 163–180). Springer Berlin Heidelberg.

Gill, D., & Prowse, V. L. (2012). Cognitive Ability and Learning to Play Equilibrium: A Level-*k* Analysis. *SSRN Electronic Journal*.

Gillan, C. M., Vaghi, M. M., Hezemans, F. H., van Ghesel Grothe, S., Dafflon, J., Brühl, A. B., Savulich, G., & Robbins, T. W. (2020). Experimentally induced and real-world anxiety have no demonstrable effect on goal-directed behaviour. *Psychological Medicine*, 1–12.

Gilovich, T., Griffin, D. W., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge University Press.

Gläscher, J. (2009). Visualization of Group Inference Data in Functional Neuroimaging. *Neuroinformatics*, *7*(1), 73–82.

Glienke, K., Wolf, O. T., & Bellebaum, C. (2015). The impact of stress on feedback and error processing during behavioral adaptation. *Neuropsychologia*, *71*, 181–190.

Grillon, C. (2008). Models and mechanisms of anxiety: Evidence from startle studies. *Psychopharmacology*, *199*(3), 421–437.

Grillon, C., Baas, J. P., Lissek, S., Smith, K., & Milstein, J. (2004). Anxious responses to predictable and unpredictable aversive events. *Behavioral Neuroscience*, *118*(5), 916–924.

Grillon, C., Duncko, R., Covington, M. F., Kopperman, L., & Kling, M. A. (2007). Acute Stress Potentiates Anxiety in Humans. *Biological Psychiatry*, *62*(10), 1183–1186.

Grillon, C., Robinson, O. J., Cornwell, B., & Ernst, M. (2019). Modeling anxiety in healthy humans: A key intermediate bridge between basic and clinical sciences. *Neuropsychopharmacology*, *44*(12), 1999–2010.

Grupe, D. W., & Nitschke, J. B. (2013). Uncertainty and anticipation in anxiety: An integrated neurobiological and psychological perspective. *Nature Reviews Neuroscience*, *14*(7), 488–501.

Grupe, D.W. (2017). Decision-Making in Anxiety and Its Disorders. In *Decision Neuroscience* (pp. 327–338). Elsevier.

Guitart-Masip, M., Huys, Q. J. M., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Go and no-go learning in reward and punishment: Interactions between affect and effect. *NeuroImage*, *62*(1), 154–166.

Haber, S. N., & Knutson, B. (2009). The Reward Circuit: Linking Primate Anatomy and Human Imaging. *Neuropsychopharmacology*, *35*(1), 4–26.

Hartley, C. A., & Phelps, E. A. (2012). Anxiety and Decision-Making. *Biological Psychiatry*, *72*(2), 113–118.

Hebart, M. N., Schriever, Y., Donner, T. H., & Haynes, J.-D. (2016). The Relationship between Perceptual Decision Variables and Confidence in the Human Brain. *Cerebral Cortex*, *26*(1), 118–130.

Heilbron, M., & Meyniel, F. (2019). Confidence resets reveal hierarchical adaptive learning in humans. *PLOS Computational Biology*, *15*(4), e1006972.

Hermans, E. J., Henckens, M. J. A. G., Joëls, M., & Fernández, G. (2014). Dynamic adaptation of large-scale brain networks in response to acute stressors. *Trends in Neurosciences*, *37*(6), 304–

Hilgenstock, R., Weiss, T., & Witte, O. W. (2014). You'd Better Think Twice: Post-Decision Perceptual Confidence. *NeuroImage*, *99*, 323–331.

Holaway, R. M., Heimberg, R. G., & Coles, M. E. (2006). A comparison of intolerance of uncertainty in analogue obsessive-compulsive disorder and generalized anxiety disorder. *Journal of Anxiety Disorders*, *20*(2), 158–174.

Hollard, G., Massoni, S., & Vergnaud, J.-C. (2016). In search of good probability assessors: An experimental comparison of elicitation rules for confidence judgments. *Theory and Decision*, *80*(3), 363–387.

Holmes, N. M., Marchand, A. R., & Coutureau, E. (2010). Pavlovian to instrumental transfer: A neurobehavioural perspective. *Neuroscience & Biobehavioral Reviews*, *34*(8), 1277–1295.

Hoven, M., Lebreton, M., Engelmann, J. B., Denys, D., Luigjes, J., & van Holst, R. J. (2019). Abnormalities of confidence in psychiatry: An overview and future perspectives. *Translational Psychiatry*, *9*(1), 268.

Hsee, C. K., & Rottenstreich, Y. (2004). Music, Pandas, and Muggers: On the Affective Psychology of Value. *Journal of Experimental Psychology: General*, *133*(1), 23–30.

Hur, J., Smith, J. F., DeYoung, K. A., Anderson, A. S., Kuang, J., Kim, H. C., Tillman, R. M., Kuhn, M., Fox, A. S., & Shackman, A. J. (2020). *Anxiety and the neurobiology of temporally uncertain threat anticipation*. 61.

Huys, Q. J. M., Browning, M., Paulus, M. P., & Frank, M. J. (2020). Advances in the computational understanding of mental illness. *Neuropsychopharmacology*.

Huys, Q. J. M., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, *19*(3), 404–413.

Jackson, E. D., Payne, J. D., Nadel, L., & Jacobs, W. J. (2006). Stress differentially modulates fear conditioning in healthy men and women. *Biological Psychiatry*, *59*(6), 516–522.

Jahfari, S., Ridderinkhof, K. R., Collins, A. G. E., Knapen, T., Waldorp, L. J., & Frank, M. J. (2019). Cross-Task Contributions of Frontobasal Ganglia Circuitry in Response Inhibition and Conflict-Induced Slowing. *Cerebral Cortex*, *29*(5), 1969–1983.

Jelen, P., Soltysik, S., & Zagrodzka, J. (2003). 22-kHz Ultrasonic vocalization in rats as an index of anxiety but not fear: Behavioral and pharmacological modulation of affective state. *Behavioural Brain Research*, *141*(1), 63–72.

Jenks, C. W., & Lawyer, S. R. (2015). Using delay discounting to understand impulsive choice in socially anxious individuals: Failure to replicate. *Journal of Behavior Therapy and Experimental Psychiatry*, *46*, 198–201.

Kahneman, D. (2011). *Thinking, fast and slow* (1st ed). Farrar, Straus and Giroux.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263.

Kahneman, D., & Tversky, A. (Eds.). (2000). *Choices, values, and frames*. Russell sage Foundation ; Cambridge University Press.

Kamstra, M. J., Kramer, L. A., & Levi, M. D. (2003). Winter Blues: A SAD Stock Market Cycle. *American Economic Review*, *93*(1), 324–343.

Karni, E. (2009). A Mechanism for Eliciting Probabilities. *Econometrica*, *77*(2), 603–606.

Kiani, R., Corthell, L., & Shadlen, M. N. (2014). Choice Certainty Is Informed by Both Evidence and Decision Time. *Neuron*, *84*(6), 1329–1342.

Kim, A. J., & Anderson, B. A. (2020). Threat reduces value-driven but not salience-driven attentional capture. *Emotion*, *20*(5), 874–889.

Kim, H., & Cabeza, R. (2009). Common and specific brain regions in high- versus low-confidence recognition memory. *Brain Research*, *1282*, 103–113.

King-Casas, B. (2005). Getting to Know You: Reputation and Trust in a Two-Person Economic Exchange. *Science*, *308*(5718), 78–83.

Kirschbaum, C., Pirke, K.-M., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test' – A Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting. *Neuropsychobiology*, *28*(1–2), 76–81.

Klein, T. A., Ullsperger, M., & Jocham, G. (2017). Learning relative values in the striatum induces violations of normative decision making. *Nature Communications*, *8*, 16033.

Klumpers, F., Kroes, M. C. W., Baas, J. M. P., & Fernández, G. (2017). How Human Amygdala and Bed Nucleus of the Stria Terminalis May Drive Distinct Defensive Responses. *The Journal of Neuroscience*, *37*(40), 9645–9656.

Knutson, B., Samanez-Larkin, G. R., & Kuhnen, C. M. (2011). Gain and Loss Learning Differentially Contribute to Life Financial Outcomes. *PLoS ONE*, *6*(9), e24390.

Koellinger, P., & Treffers, T. (2015). Joy Leads to Overconfidence, and a Simple Countermeasure. *PLOS ONE*, *10*(12), e0143263.

Konovalov, A., & Krajbich, I. (2019). Over a Decade of Neuroeconomics: What Have We Learned? *Organizational Research Methods*, *22*(1), 148–173.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience*, *12*(5), 535–540.

Lebreton, M., Abitbol, R., Daunizeau, J., & Pessiglione, M. (2015). Automatic integration of confidence in the brain valuation signal. *Nature Neuroscience*, *18*(8), 1159–1167.

Lebreton, M., Bacily, K., Palminteri, S., & Engelmann, J. (2019). *Contextual influence on confidence judgments in human reinforcement learning*. *4*(e1006973), 27.

Lebreton, M., Bacily, K., Palminteri, S., & Engelmann, J. B. (2018). Contextual influence on confidence judgments in human reinforcement learning. *BioRxiv*, 339382.

Lebreton, M, Langdon, S., Slieker, M. J., Nooitgedacht, J. S., Goudriaan, A. E., Denys, D., van Holst, R. J., & Luigjes, J. (2018). Two sides of the same coin: Monetary incentives concurrently improve and bias confidence judgments. *SCIENCE ADVANCES*, 14.

LeDoux, J. E., & Pine, D. S. (2016). Using Neuroscience to Help Understand Fear and Anxiety: A Two-System Framework. *American Journal of Psychiatry*, *173*(11), 1083–1093.

Lee, D., Seo, H., & Jung, M. W. (2012). Neural basis of reinforcement learning and decision making. *Annual Review of Neuroscience*, *35*, 287–308.

Lefebvre, G., Lebreton, M., Meyniel, F., Bourgeois-Gironde, S., & Palminteri, S. (2017). Behavioural and neural characterization of optimistic reinforcement learning. *Nature Human Behaviour*, *1*(4).

Lempert, K. M., Glimcher, P. W., & Phelps, E. A. (2015). Emotional arousal and discount rate in intertemporal choice are reference dependent. *Journal of Experimental Psychology: General*, *144*(2), 366–373.

Lempert, K. M., Porcelli, A. J., Delgado, M. R., & Tricomi, E. (2012). Individual Differences in Delay Discounting Under Acute Stress: The Role of Trait Perceived Stress. *Frontiers in Psychology*, *3*.

Lep, Ž., Babnik, K., & Hacin Beyazoglu, K. (2020). Emotional Responses and Self-Protective Behavior Within Days of the COVID-19 Outbreak: The Promoting Role of Information Credibility. *Frontiers in Psychology*, *11*, 1846.

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and Decision Making. *Annual Review of Psychology*, *66*(1), 799–823.

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: A neural common currency for choice. *Current Opinion in Neurobiology*, *22*(6), 1027–1038.

Lighthall, N. R., Gorlick, M. A., Schoeke, A., Frank, M. J., & Mather, M. (2013). Stress modulates reinforcement learning in younger and older adults. *Psychology and Aging*, *28*(1), 35–46.

Lindström, B., Selbing, I., Molapour, T., & Olsson, A. (2014). Racial Bias Shapes Social Reinforcement Learning. *Psychological Science*, *25*(3), 711–719.

Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, *28*(2), 129–137.

Loewenstein, G. (2000). Emotions in Economic Theory and Economic Behavior. *American Economic Review*, *90*(2), 426–432.

Lopez-Persem, A., Bastin, J., Petton, M., Abitbol, R., Lehongre, K., Adam, C., Navarro, V., Rheims, S., Kahane, P., Domenech, P., & Pessiglione, M. (2020). Four core properties of the human brain valuation system demonstrated in intracranial signals. *Nature Neuroscience*, *23*(5), 664–675.

Louie, K., Glimcher, P. W., & Webb, R. (2015). Adaptive neural coding: From biological to behavioral decision-making. *Current Opinion in Behavioral Sciences*, *5*, 91–99.

Mahlberg, J., Seabrooke, T., Weidemann, G., Hogarth, L., Mitchell, C. J., & Moustafa, A. A. (2019). Human appetitive Pavlovian-to-instrumental transfer: A goal-directed account. *Psychological Research*.

Malmendier, U., & Tate, G. (2005). CEO Overconfidence and Corporate Investment. *The Journal of Finance*, *60*(6), 2661–2700.

Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, *21*(1), 422–430.

Massoni, S. (2014). Emotion as a boost to metacognition: How worry enhances the quality of confidence. *Consciousness and Cognition*, *29*, 189–198.

Mather, M., & Lighthall, N. R. (2012). Risk and Reward Are Processed Differently in Decisions Made Under Stress. *Current Directions in Psychological Science*, *21*(1), 36–41.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315.

Meier, A. N., Schmid, L., & Stutzer, A. (2019). Rain, emotions and voting for the status quo. *European Economic Review*, *119*, 434–451.

Meyniel, F., Sigman, M., & Mainen, Z. F. (2015). Confidence as Bayesian Probability: From Neural Origins to Behavior. *Neuron*, *88*(1), 78–92.

Miele, D. B., Wager, T. D., Mitchell, J. P., & Metcalfe, J. (2011). Dissociating Neural Correlates of Action Monitoring and Metacognition of Agency. *Journal of Cognitive Neuroscience*, *23*(11), 3620–3636.

Mkrtchian, A., Aylward, J., Dayan, P., Roiser, J. P., & Robinson, O. J. (2017). Modeling Avoidance in Mood and Anxiety Disorders Using Reinforcement Learning. *Biological Psychiatry*, *82*(7), 532–539.

Molenberghs, P., Trautwein, F.-M., Böckler, A., Singer, T., & Kanske, P. (2016). Neural correlates of metacognitive ability and of feeling confident: A large-scale fMRI study. *Social Cognitive and Affective Neuroscience*, *11*(12), 1942–1951.

Montague, P. R., & Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron*, *36*(2), 265-284.

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences*, *16*(1), 72–80.

Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature*, *431*(7010), 760–767.

Moran, R., Teodorescu, A. R., & Usher, M. (2015). Post choice information integration as a causal determinant of confidence: Novel data and a computational account. *Cognitive Psychology*, *78*, 99–147.

Moran, T. P. (2016). Anxiety and working memory capacity: A meta-analysis and narrative review. *Psychological Bulletin*, *142*(8), 831–864.

Mowrer, O. H. (1952). Chapter V: Learning Theory. *Review of Educational Research*, *22*(5), 475–495.

Mukherjee, K. (2010). A dual system model of preferences under risk. *Psychological Review*, *117*(1), 243–255.

Nassar, M. R., & Gold, J. I. (2013). A Healthy Fear of the Unknown: Perspectives on the Interpretation of Parameter Fits from Computational Models in Neuroscience. *PLoS Computational Biology*, *9*(4), e1003015.

Navajas, J., Bahrami, B., & Latham, P. E. (2016). Post-decisional accounts of biases in confidence. *Current Opinion in Behavioral Sciences*, *11*, 55–60.

Nieuwenhuis, S., Heslenfeld, D. J., Alting von Geusau, N. J., Mars, R. B., Holroyd, C. B., & Yeung, N. (2005). Activity in human reward-sensitive brain areas is strongly context dependent. *NeuroImage*, *25*(4), 1302–1309.

Oldham, S., Murawski, C., Fornito, A., Youssef, G., Yücel, M., & Lorenzetti, V. (2018). The anticipation and outcome phases of reward and loss processing: A neuroimaging meta-analysis of the monetary incentive delay task. *Human Brain Mapping*, *39*(8), 3398–3418.

Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A., & Daw, N. D. (2013). Working-memory capacity protects model-based learning from stress. *Proceedings of the National Academy of Sciences*, *110*(52), 20941–20946.

Padoa-Schioppa, C. (2009). Range-Adapting Representation of Economic Value in the Orbitofrontal Cortex. *Journal of Neuroscience*, *29*(44), 14004–14014.

Palminteri, S., & Chevallier, C. (2018). Can We Infer Inter-Individual Differences in Risk-Taking From Behavioral Tasks? *Frontiers in Psychology*, *9*, 2307.

Palminteri, S., Justo, D., Jauffret, C., Pavlicek, B., Dauta, A., Delmaire, C., Czernecki, V., Karachi, C., Capelle, L., & Durr, A. (2012). Critical roles for anterior insula and dorsal striatum in punishment-based avoidance learning. *Neuron*, *76*(5), 998–1009.

Palminteri, S., Justo, D., Jauffret, C., Pavlicek, B., Dauta, A., Delmaire, C., Czernecki, V., Karachi, C., Capelle, L., Durr, A., & Pessiglione, M. (2012). Critical Roles for Anterior Insula and Dorsal Striatum in Punishment-Based Avoidance Learning. *Neuron*, *76*(5), 998–1009.

Palminteri, S., Khamassi, M., Joffily, M., & Coricelli, G. (2015). Contextual modulation of value signals in reward and punishment learning. *Nature Communications*, *6*(1).

Palminteri, S., Kilford, E. J., Coricelli, G., & Blakemore, S.-J. (2016). The Computational Development of Reinforcement Learning during Adolescence. *PLOS Computational Biology*, *12*(6), e1004953.

Palminteri, S., Lefebvre, G., Kilford, E. J., & Blakemore, S.-J. (2017). Confirmation bias in human reinforcement learning: Evidence from counterfactual feedback processing. *PLoS Computational Biology*, *13*(8), e1005684.

Palminteri, S., & Pessiglione, M. (2017). Opponent Brain Systems for Reward and Punishment Learning. In *Decision Neuroscience* (pp. 291–303). Elsevier.

Palminteri, S., Wyart, V., & Koechlin, E. (2017). The importance of falsification in computational cognitive modeling. *Trends in Cognitive Sciences*, *21*(6), 425–433. h

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*(2), 220–244. h

Perlstein, W. M., Elbert, T., & Stenger, V. A. (2002). Dissociation in human prefrontal cortex of affective influences on working memory-related activity. *Proceedings of the National Academy of Sciences*, *99*(3), 1736–1741.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*(7106), 1042–1045.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, *9*(2), 148–158.

Pessoa, L., & Engelmann, J. B. (2010). Embedding Reward Signals into Perception and Cognition. *Frontiers in Neuroscience*, *4*.

Peters, E., Västfjäll, D., Gärling, T., & Slovic, P. (2006). Affect and decision making: A "hot" topic. *Journal of Behavioral Decision Making*, *19*(2), 79–85.

Petzold, A., Plessow, F., Goschke, T., & Kirschbaum, C. (2010). Stress reduces use of negative feedback in a feedback-based learning task. *Behavioral Neuroscience*, *124*(2), 248–255.

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the Amygdala to Emotion Processing: From Animal Models to Human Behavior. *Neuron*, *48*(2), 175–187.

Phelps, E. A., Lempert, K. M., & Sokol-Hessner, P. (2014). Emotion and Decision Making: Multiple Modulatory Neural Circuits. *Annual Review of Neuroscience*, *37*(1), 263–287.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS*. Springer.

Plassmann, H., O'Doherty, J., Shiv, B., & Rangel, A. (2008). Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of the National Academy of Sciences*, *105*(3), 1050–1054.

Pleskac, T. J., & Busemeyer, J. (2007). *A Dynamic and Stochastic Theory of Choice, Response Time, and Confidence*. 7.

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: A theory of choice, decision time, and confidence. *Psychological Review*, *117*(3), 864–901.

Porcelli, A. J., & Delgado, M. R. (2017). Stress and decision making: Effects on valuation, learning, and risk-taking. *Current Opinion in Behavioral Sciences*, *14*, 33–39.

Porcelli, A. J., Lewis, A. H., & Delgado, M. R. (2012). Acute stress influences neural circuits of reward processing. *Frontiers in Neuroscience*, *6*.

Pouget, A., Drugowitsch, J., & Kepecs, A. (2016). Confidence and certainty: Distinct probabilistic quantities for different goals. *Nature Neuroscience*, *19*(3), 366–374.

Pulcu, E., & Browning, M. (2019). The Misestimation of Uncertainty in Affective Disorders. *Trends in Cognitive Sciences*, *23*(10), 865–875.

Qiu, L., Su, J., Ni, Y., Bai, Y., Zhang, X., Li, X., & Wan, X. (2018). The neural system of metacognition accompanying decision-making in the prefrontal cortex. *PLOS Biology*, *16*(4), e2004037.

Raio, C. M., Hartley, C. A., Orederu, T. A., Li, J., & Phelps, E. A. (2017). Stress attenuates the flexible updating of aversive value. *Proceedings of the National Academy of Sciences*, *114*(42), 11241–11246.

Rangel, A. (2009). The Computation and Comparison of Value in Goal-directed Choice. In *Neuroeconomics* (pp. 425–440). Elsevier.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545–556.

Ratcliff, R., & Starns, J. J. (2009). Modeling confidence and response time in recognition memory. *Psychological Review*, *116*(1), 59–83.

Ratcliff, R., & Starns, J. J. (2013). Modeling confidence judgments, response times, and multiple choices in decision making: Recognition memory and motion discrimination. *Psychological Review*, *120*(3), 697–719.

Rescorla, R. A., & Solomon, R. L. (1967). Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psychological Review*, *74*(3), 151.

Rescorla, R. A., & Wagner, A. R. (1972). *A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement*. *2*, 64–99.

Robinson, O. J., Bond, R. L., & Roiser, J. P. (2015a). The impact of stress on financial decision-making varies as a function of depression and anxiety symptoms. *PeerJ*, *3*, e770.

Robinson, O. J., Bond, R. L., & Roiser, J. P. (2015b). The impact of threat of shock on the framing effect and temporal discounting: Executive functions unperturbed by acute stress? *Frontiers in Psychology*, *6*.

Robinson, O. J., & Chase, H. W. (2017). Learning and Choice in Mood Disorders: Searching for the Computational Parameters of Anhedonia. *Computational Psychiatry*, *1*, 208–233.

Robinson, O. J., Overstreet, C., Charney, D. R., Vytal, K., & Grillon, C. (2013). Stress increases aversive prediction error signal in the ventral striatum. *Proceedings of the National Academy of Sciences*, *110*(10), 4129–4133.

Robinson, O. J., Vytal, K., Cornwell, B. R., & Grillon, C. (2013). The impact of anxiety upon cognition: Perspectives from human threat of shock studies. *Frontiers in Human Neuroscience*, *7*, 1:21.

Roelofs, K. (2017). Freeze for action: Neurobiological mechanisms in animal and human freezing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1718), 20160206.

Roelofs, K., Hagenaars, M. A., & Stins, J. (2010). Facing Freeze: Social Threat Induces Bodily Freeze in Humans. *Psychological Science*, *21*(11), 1575–1581.

Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, *8*(1), 164–212.

Rottenstreich, Y., & Hsee, C. K. (2001). Money, Kisses, And Electric Shocks. *Psychological Science*, *12*(3), 6.

Rouault, M., Seow, T., Gillan, C. M., & Fleming, S. M. (2018). Psychiatric Symptom Dimensions Are Associated With Dissociable Shifts in Metacognition but Not Task Performance. *Biological Psychiatry*, *84*(6), 443–451.

Rounds, J. S., Beck, J. G., & Grant, D. M. (2007). Is the delay discounting paradigm useful in understanding social anxiety? *Behaviour Research and Therapy*, *45*(4), 729–735.

Rounis, E., Maniscalco, B., Rothwell, J. C., Passingham, R. E., & Lau, H. (2010). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cognitive Neuroscience*, *1*(3), 165–175.

Rutledge, R. B., & Adams, R. A. (2017). Computational Psychiatry. In A. A. Moustafa (Ed.), *Computational Models of Brain and Behavior* (pp. 29–42). John Wiley & Sons, Ltd.

Sadeghi, S., Ekhtiari, H., Bahrami, B., & Ahmadabadi, M. N. (2017). Metacognitive Deficiency in a Perceptual but Not a Memory Task in Methadone Maintenance Patients. *Scientific Reports*, *7*(1), 7052.

Safra, L., Chevallier, C., & Palminteri, S. (2018). Social information impairs reward learning in depressive subjects: Behavioral and computational characterization. *BioRxiv*.

Sanders JI, Hangya B, Kepecs A. Signatures of a Statistical Computation in the Human Sense of Confidence. Neuron. 2016 May 4;90(3):499-506.

Savitsky, K., Medvec, V. H., Charlton, A. E., & Gilovich, T. (1998). "What, Me Worry?": Arousal, Misattribution, and the Effect of Temporal Distance on Confidence. *Personality and Social Psychology Bulletin*, *24*(5), 529–536.

Schindler, S., & Pfattheicher, S. (2017). The frame of the game: Loss-framing increases dishonest behavior. *Journal of Experimental Social Psychology*, *69*, 172–177.

Schlag, K. H., Tremewan, J., & van der Weele, J. J. (2015). A penny for your thoughts: A survey of methods for eliciting beliefs. *Experimental Economics*, *18*(3), 457–490.

Schmitz, A., & Grillon, C. (2012). Assessing fear and anxiety in humans using the threat of predictable and unpredictable aversive events (the NPU-threat test). *Nature Protocols*, *7*(3), 527–532.

Schonberg, T., Fox, C. R., & Poldrack, R. A. (2011). Mind the gap: Bridging economic and naturalistic risk-taking with cognitive neuroscience. *Trends in Cognitive Sciences*, *15*(1), 11–19.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599.

Schwabe, L., & Wolf, O. T. (2009). Stress prompts habit behavior in humans. *Journal of Neuroscience*, *29*(22), 7191–7198.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*(3), 513–523.

Seymour, B., Maruyama, M., & De Martino, B. (2015). When is a loss a loss? Excitatory and inhibitory processes in loss-related decision-making. *Current Opinion in Behavioral Sciences*, *5*, 122–127.

Seymour, B., & McClure, S. M. (2008). Anchors, scales and the relative coding of value in the brain. *Current Opinion in Neurobiology*, *18*(2), 173–178.

Shapiro, A. D., & Grafton, S. T. (2020). Subjective value then confidence in human ventromedial prefrontal cortex. *PLOS ONE*, *15*(2), e0225617.

Shenhav, A., Straccia, M. A., Cohen, J. D., & Botvinick, M. M. (2014). Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nature Neuroscience*, *17*(9), 1249–1254.

Sidi, Y., Ackerman, R., & Erez, A. (2018). Feeling happy and (over)confident: The role of positive affect in metacognitive processes. *Cognition and Emotion*, *32*(4), 876–884.

Sip, K. E., Muratore, A. F., & Stern, E. R. (2016). Effects of context on risk taking and decision times in obsessive-compulsive disorder. *Journal of Psychiatric Research*, *75*, 82–90.

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European Journal of Operational Research*, *177*(3), 1333–1352.

Sokol-Hessner, P., Hsu, M., Curley, N. G., Delgado, M. R., Camerer, C. F., & Phelps, E. A. (2009). Thinking like a trader selectively reduces individuals' loss aversion. *Proceedings of the National Academy of Sciences*, *106*(13), 5035–5040.

Sokol-Hessner, P., & Rutledge, R. B. (2019). *The Psychological and Neural Basis of Loss Aversion*.

Spielberger, C., Gorsuch, R., Lushene, R., Vagg, P., & Jacobs, G. (1983). *Manual for the state-trait anxiety inventory*. Palo Alto, CA: Consulting Psychologists.

Steinglass, J. E., Lempert, K. M., Choo, T.-H., Kimeldorf, M. B., Wall, M., Walsh, B. T., Fyer, A. J., Schneier, F. R., & Simpson, H. B. (2017). Temporal discounting across three psychiatric disorders: Anorexia nervosa, obsessive compulsive disorder, and social anxiety disorder. *Depression and Anxiety*, *34*(5), 463–470.

Stevens, S., Peters, A., Abraham, A., & Hermann, C. (2014). Enhanced avoidance behavior in social anxiety: Evidence from a probabilistic learning task. *Journal of Behavior Therapy and Experimental Psychiatry*, *45*(1), 39–45.

Story, G. W., Vlaev, I., Seymour, B., Winston, J. S., Darzi, A., & Dolan, R. J. (2013). Dread and the Disvalue of Future Pain. *PLoS Computational Biology*, *9*(11), e1003335.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. MIT Press.

Sylvers, P., Lilienfeld, S. O., & LaPrairie, J. L. (2011). Differences between trait fear and trait anxiety: Implications for psychopathology. *Clinical Psychology Review*, *31*(1), 122–137.

Talmi, D., Dayan, P., Kiebel, S. J., Frith, C. D., & Dolan, R. J. (2009). How Humans Integrate the Prospects of Pain and Reward during Choice. *Journal of Neuroscience*, *29*(46), 14617–14626.

Talmi, Deborah, Seymour, B., Dayan, P., & Dolan, R. J. (2008). Human Pavlovian–Instrumental Transfer. *Journal of Neuroscience*, *28*(2), 360–368.

Thaler, R. H., & Johnson, E. J. (1990). Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice. *Management Science*, *36*(6), 643–660.

Ting, C.-C., Palminteri, S., Lebreton, M., & Engelmann, J. B. (2020a). *The elusive effects of incidental anxiety on reinforcement-learning* [Preprint]. PsyArXiv.

Ting, C.-C., Palminteri, S., Engelmann, J. B., & Lebreton, M. (2020b). Robust valence-induced biases on motor response and confidence in human reinforcement learning. *Cognitive, Affective, & Behavioral Neuroscience*.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The Neural Basis of Loss Aversion in Decision-Making Under Risk. *Science*, *315*(5811), 515–518.

Trapp, S., O'Doherty, J. P., & Schwabe, L. (2018). Stressful events as teaching signals for the brain. *Trends in Cognitive Sciences*, *22*(6), 475–478.

Treadway, M. T., Admon, R., Arulpragasam, A. R., Mehta, M., Douglas, S., Vitaliano, G., Olson, D. P., Cooper, J. A., & Pizzagalli, D. A. (2017). Association between interleukin-6 and striatal prediction-error signals following acute stress in healthy female participants. *Biological Psychiatry*, *82*(8), 570–577.

Tymula, A., & Plassmann, H. (2016). Context-dependency in valuation. *Current Opinion in Neurobiology*, *40*, 59–65.

van den Berg, R., Anandalingam, K., Zylberberg, A., Kiani, R., Shadlen, M. N., & Wolpert, D. M. (2016). A common mechanism underlies changes of mind about decisions and confidence. *ELife*, *5*, e12192.

van den Bos, W., Talwar, A., & McClure, S. M. (2013). Neural Correlates of Reinforcement Learning and Social Preferences in Competitive Bidding. *Journal of Neuroscience*, *33*(5), 2137–2146.

Vickers, D., Smith, P., Burt, J., & Brown, M. (1985). Experimental paradigms emphasising state or process limitations: II effects on confidence. *Acta Psychologica*, *59*(2), 163–193.

Vinckier, F., Gaillard, R., Palminteri, S., Rigoux, L., Salvador, A., Fornito, A., Adapa, R., Krebs, M. O., Pessiglione, M., & Fletcher, P. C. (2016). Confidence and psychosis: A neuro-computational account of contingency learning disruption by NMDA blockade. *Molecular Psychiatry*, *21*(7), 946–955.

Voegler, R., Peterburs, J., Bellebaum, C., & Straube, T. (2019). Modulation of feedback processing by social context in social anxiety disorder (SAD)–an event-related potentials (ERPs) study. *Scientific Reports*, *9*(1).

Vytal, K. E., Cornwell, B. R., Letkiewicz, A. M., Arkin, N. E., & Grillon, C. (2013). The complex interaction between anxiety and cognition: Insight from spatial and verbal working memory. *Frontiers in Human Neuroscience*, *7*.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070.

Watson, P., Wiers, R. W., Hommel, B., & de Wit, S. (2014). Working for food you don't desire. Cues interfere with goal-directed food-seeking. *Appetite*, *79*, 139–148.

Williams, L. M., Brown, K. J., Das, P., Boucsein, W., Sokolov, E. N., Brammer, M. J., Olivieri, G., Peduto, A., & Gordon, E. (2004). The dynamics of cortico-amygdala and autonomic activity over the experimental time course of fear perception. *Cognitive Brain Research*, *21*(1), 114–123.

Wilson, R. C., & Collins, A. G. (2019). Ten simple rules for the computational modeling of behavioral data. *ELife*, *8*, e49547.

Wisco, B. E., Marx, B. P., Wolf, E. J., Miller, M. W., Southwick, S. M., & Pietrzak, R. H. (2014). Posttraumatic Stress Disorder in the US Veteran Population: Results From the National Health and Resilience in Veterans Study. *The Journal of Clinical Psychiatry*, *75*(12), 1338–1346.

Wittmann, M. K., Kolling, N., Faber, N. S., Scholl, J., Nelissen, N., & Rushworth, M. F. S. (2016). Self-Other Mergence in the Frontal Cortex during Cooperation and Competition. *Neuron*, *91*(2), 482–493.

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: Confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1310–1321.

Yu, S., Pleskac, T. J., & Zeigenfuse, M. D. (2015). Dynamics of postdecisional processing of confidence. *Journal of Experimental Psychology: General*, *144*(2), 489–510.

Zenger, M., Glaesmer, H., Hockel, M., & Hinz, A. (2011). Pessimism Predicts Anxiety, Depression and Quality of Life in Female Cancer Patients. *Japanese Journal of Clinical Oncology*, *41*(1), 87–94.

Nederlandse Samenvatting

Het effect van contexten op besluitvorming is zichtbaar in veld en laboratorium experimenten. De mechanismes van context geïnduceerde veranderingen op keuze patronen zijn breed onderzocht op zowel gedrag als neuraal niveau. Bijvoorbeeld, referentiepunten in vooruitzichttheorie – beter bekend als "prospect theory" – verklaren waarom de waarde van de zelfde set aan opties als ongelijk wordt berekend. Echter, het is minder duidelijke hoe affect en voorgaande prestaties zijn betrokken bij besluitvorming na verloop van tijd. "Reinforcement learning" biedt een kader om dit probleem aan te pakken door dit updateproces te kwantificeren. Het huidige proefschrift combineert een gemodificeerde reinforcement learning taak met werkwijzen van affectieve neurowetenschap, computationele modellen en beeldvorming van de hersenen om op systematische wijze de onderliggende mechanismes van het effect van twee affectieve context manipulaties (d.w.z., keuze- irrelevante affect: incidentele angst; keuze-relevante affect: uitkomstvalentie) op het bijwerken van overtuigingen (d.w.z., optiewaarde en zekerheidsoordelen) te onderzoeken. Onze resultaten demonstreren dat incidentele angst en uitkomstvalentie een beperkt effect hebben op leren an sich. Belangrijker, het effect van uitkomstvalentie op reactie tijd en zekerheidsoordelen is robuust. Deze bevindingen impliceren niet alleen dat zekerheid, optiewaarde en context tegelijkertijd worden bijgewerkt, maar de robuuste valentie-geïnduceerde veranderingen in zekerheid duiden aan dat metacognitieve werking ontvankelijker is voor affectieve context manipulaties. Behalve het empirische bewijs met betrekking tot de rol van angst en uitkomstvalentie in reinforcement learning, hebben wij ook nuttige methodologische en analytische aanpakken opgesteld voor toekomstige studies die onderzoek doen naar de interacties tussen emotie, cognitie en metacognitie in het veld van oordeel- en besluitvorming. De belangrijkste bevindingen voor elk hoofdstuk zijn als volgt afzonderlijk samengevat:

Hoofdstuk 2 onderzoekt hoe en of incidentele angst instrumenteel leren beïnvloedt, terwijl tegelijkertijd verscheidene problemen in eerder onderzoek worden aangekaart en verbeterd door middel van een gericht literatuuronderzoek. Wij hebben een krachtig binnen-proefpersonen design gebruikt met zowel een leer- als overdrachtsfase en 2 affectieve context manipulaties: omgeving (angst versus veilig) en uitkomstvalentie (winst versus verlies). In twee varianten (N = 2x50) van deze experimentele opzet is incidentele angst geïnduceerd door onverwachte, aversieve en prestatie-onafhankelijke elektrische schokken toe te dienen tijdens de leertaak. Daarbij is de veroorzaakte angst beoordeeld door middel van zowel een vragenlijst als een fysiologische respons (d.w.z., huidgeleidingsrespons). Een uitgebreide modelleringsinspanning onthulde dat, onafhankelijk van de effecten van angst, individuen zwaarder tillen aan positieve dan negatieve uitkomsten en de neiging hebben om een omissie van verlies als winst te ervaren (en omgekeerd). Echter, in lijn met de resultaten van ons gericht literatuuronderzoek bleek het isoleren van de specifieke computationele effecten van angst op leren an sich een uitdaging te zijn. Algeheel suggereren onze resultaten dat leermechanismen complexer zijn dan traditioneel verondersteld en roepen ze belangrijke vragen op over de robuustheid van de eerder geïdentificeerde effecten van angst in eenvoudige reinforcement learning studies.

Hoofdstuk 3 focust op de impact van keuze gerelateerd affect (d.w.z., uitkomstvalentie: winst versus verlies) op 3 veelvoorkomende maten: nauwkeurigheid, zekerheid en reactietijd (RT). Hoewel deze drie maten sterk gecorreleerd zouden zijn, blijft het controversieel of valentie-geïnduceerde zekerheid en RT veranderingen van elkaar te onderscheiden zijn. Om dit vraagstuk aan te pakken, is het doel van hoofdstuk 3 om de aanwezigheid van de valentie-geïnduceerde zekerheid tendens vast te stellen in de afwezigheid van de RT tendens. Wij hebben zes varianten van een leertaak uitgevoerd, en hebben geprobeerd om de valentie-geïnduceerde motor tendens effecten te ontwrichten door de arrangering tussen keuzes en acties te manipuleren en beperkingen op te leggen op de respons tijden (RTs). Wij hebben zowel motor als zekerheid tendensen waargenomen ondanks onze ontkoppelingspogingen, vaststellend dat de effecten van valentie op motor en metacognitieve responsen zeer robuust en repliceerbaar zijn. Desondanks onthullen gevolgtrekkingen binnen en tussen individuen dat de zekerheidstendens de ontkoppeling van de RT tendens weerstaat. Derhalve, hoewel geassocieerd in de meeste gevallen, lijken valentie-geïnduceerde motor en zekerheid tendensen gedeeltelijk dissocieerbaar. Deze resultaten benadrukken nieuwe belangrijke mechanistische beperkingen die opgenomen moeten worden in leermodellen om gezamenlijk keuze, reactie tijden en zekerheid te verklaren.

Hoofdstuk 4 toetst of de vorming van zekerheid in het brein contextafhankelijk of contextonafhankelijk is. Wij hebben fMRI en een reinforcement learning taak, geoptimaliseerd in hoofdstuk 3, gecombineerd, welke winst en verlies contexten dissocieerde en motor responsen van optie evaluatie isoleerde. Deze combinatie stelde ons in staat om taak-gerelateerde hersenactiviteit te meten en de hersengebieden betrokken bij het verwerken van zekerheid verder te identificeren. In deze fMRI studie (N=40), hebben wij succesvol het effect van valentie op zekerheid gerepliceerd, en valentie-geïnduceerd zekerheid toenames en valentie-geïnduceerde RT vertragingen gedissocieerd. Bovendien vonden we dat het waardesysteem van de hersenen in het algemeen betrokken was bij het coderen van zekerheid tijdens de presentatie van symbolen en geen enkele andere hersengebieden codeerde specifiek zekerheidsoordelen in winst en verlies contexten. Desondanks werd consequent de correlatie tussen gerapporteerde zekerheid en activatie in deze gebieden enkel gevonden in winst contexten en niet in verlies contexten. Deze resultaten tonen aan dat (1) zekerheid gevormd en gecodeerd werd in algemene neurale netwerken voor expliciete besluitvorming, en (2) gerapporteerde zekerheid in verlies contexten misschien niet compleet wordt verwerkt tijdens optie evaluatie.

Acknowledgements

Learning takes time and the chapters in this thesis took around 4.5 years. I've been fortunate to have had the opportunity to learn and to work with very pleasant and helpful people in a stimulating working environment. I am very grateful to all of them for the help that they afforded me in whatever form. I can not complete my PhD thesis without these advises/helps to a new arrival in the Netherlands, a PhD student, a new parent and finally in the defense of my thesis. From the time before my PhD journey, I was also blessed with outstanding help. I thank my current and previous supervisors and Erdem for encouraging me to step out of my comfort zone and giving me a chance to study human decision-making. The new journey of my life kicked off in 2016 in the Netherlands with two new roles: as a Ph.D student and a husband. Such a journey is truly an adventure. During this time I not only learned how to perform new roles, but also I adapted to a new culture, a new social system and a new lifestyle. Some experiences were challenging and even harsh but I'm pleased to say that I was able to rise to the challenges and I learned a lot about myself with supports from family, secretaries, colleagues and friends.

In particular, I would like to express my immense gratitude to my supervisors: Jan and Maël, for their professional advices, patience and insights. Jan, your unstinting encouragement and feedback always made me more confident to explore new approaches and to try a variety of opportunities, including writing a journal club article, the summer/winter school, the scholarship and the post-doctorate applications. I also appreciate the flexibility you gave me on setting up the facilities (e.g., shock device) and courses preparation. Your faith in me has been the best motivation to becoming an independent researcher. In addition, your suggestions when I hit a brick wall during my teaching and research phases were invaluable. Your good humour and concern for my well-being and career development were always a source of comfort and are much appreciated. I would now like to express my gratitude to Maël for his patient guidance. Working with you was a priceless experience because you clearly demonstrated the concept of "*less is more*" on many occasions including project management, data analysis, presentation and most importantly, research attitude. Another good lesson you taught me was the value of experimental design and the value of behavioral experiments. To be honest, these experiences were not always palatable, but I am sure good working habits and attitude will stand me in good stead for life. In addition to Jan and Maël, I would like to thank Stefano, the main co-author of our projects, for many inspiring papers and advice on data analyses and revisions. There is no doubt that our projects were gratefully improved by your views, findings and paradigms.

I would also like to thank the Master's students I supervised. Their contributions, independence and self-reliance stimulated my thinking and facilitated the speed and quality of the experiments. I would like to especially thank two students who worked with me at the very beginning of two projects. Isabela, thanks for helping out on setting up the lab and creating the protocol for the ToS projects. In addition to the lab work, I gained many insights and saw the bright side of the events when I talked to you about my pressure at work. Tiffany, it was a great experience to

conduct fMRI project with you at UvA. We prepared almost all materials together, including the preparation for the training exam. With your assistance, we successfully gathered fMRI data within 1.5 months, which was amazing.

I gratefully appreciate two groups of people who enriched my PhD journey. Members in the Affective Neuroeconomics Lab: Alejandro, Ava, Dianna, Femka, Isabela, Jan (Engelmann), Jan (Hausfeld), Joaquina, Judy, Karin, Kassandra, Konstantinos, Li-Ang, Lotte, Maël, Manon, O'Jay, Olivia, Ruth and Tiffany. We enjoyed snacks (and sometimes lunch) and science in the lab meetings. Importantly, many issues were brainstormed and solved via idea and experience exchanges during or even after the meetings. I benefited from the lab most I think, because I also practiced my presentations (including the recent one for the job interview) several times and obtained many constructive feedbacks from you all. I would like to thank colleagues at CREED: Ailko, Alejandro, Aljaž, Arthur, Aslihan, Andreas, Ayse, Carsten, Charlotte, Chris, Davide, Dianna, Frans, Giorgia, Ivan, Jan (Engelmann), Jan (Hausfeld), Joël, Joep, Jeroen, Jindi, Junze, Johan, Katharina, Kostas, Li-Ang, Margarita, Natalie, Nils, Sander, Shaul, Stephan, Theo, and Vadim. The combinations of CREED's (in)formal events and CREEDers provided me with warm and great research environments. These events and individuals encouraged me to switch between my working and relaxing mode rather than "working mode only", which is indeed a good practice for me to increase productivity.

A special thanks goes to Rasheed. Thank you for your helpful language and history lessons. I miss the time we spent on language exchange in the park, library and eventually in your apartment. It is hard to count how many weekends we spent together chatting and practising. I am convinced that you in particular enhanced my speaking and writing skills and my appreciation of history.

Thanks to Anil, Davide, Dianna, Jan, Maël, Rasheed, Stefano and many others, for giving me advices on thesis writing. I wouldn't have been able to complete and improve my thesis without your feedback. I am also very grateful for the financial support from the Amsterdam School of Economic and the Taiwan scholarship program (Government Scholarship to Study Abroad /GSSA).

Last but not least, I owe a great deal to my family. My greatest role-model has always been my dear Mom. Thank you Mom for supporting my studies abroad and sorry for letting you face so many challenges alone in the last four years, not least looking after Papa and looking for a new apartment. A final word of appreciation for my devoted and adorable wife Hsin-Yu. Thank you for sharing domestic and professional concerns and thank you for creating a wonderful home for our family which has enabled me to develop a healthier 'work/life' balance. Hsuan-Yu, my dear son. Thank you for coming into our lives in 2019. In spite of your naughtiness your smile always melts my heart. Time flies, you are almost two and you are developing into a fine young boy. I'm proud of you and thankful to your mother for her devotion.

Amsterdam, 2021