Modeling and Analysis of Affective Influences on Human Experience, Prediction, Decision Making, and Behavior

by

Hyung-il Ahn

S.M., Massachusetts Institute of Technology (2004)
B.S., Seoul National University (1998)

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
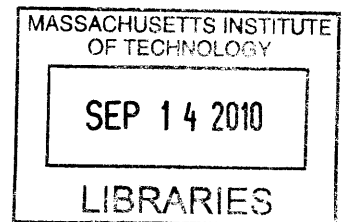in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Media Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2010

Author_____
Program in Media Arts and Sciences
August 6, 2010

Certified by_____
Rosalind W. Picard
Professor of Media Arts and Sciences
Program in Media Arts and Sciences, MIT
Thesis Supervisor

Accepted by_____
Pattie Maes
Associate Academic Head
Program in Media Arts and Sciences, MIT

Modeling and Analysis of Affective Influences on Human Experience, Prediction, Decision Making, and Behavior

by

Hyung-il Ahn

Abstract

Subjective and affective elements are well-known to influence human decision making. This dissertation presents a theoretical and empirical framework on how human decision makers' subjective experience and affective prediction influence their choice behavior under uncertainty, frames and emotions. The framework extends and integrates existing theories of prospect theory (PT) and reinforcement learning (RL), drawing on a growing literature offering the role of affect in decision making and the neural underpinnings of human decision behavior. The proposed Affective-Cognitive (AC) model extends Prospect Theory (PT)-based subjective value functions to model human experienced-utility and predicted-utility functions. The AC model assumes that the shapes (or parameters) of these subjective value functions dynamically vary with the decision makers affective states in sequential decision making. Human decision-making experiments were conducted to empirically infer how people adjust the parameters (i.e., shape and reference point) of their experienced-utility and predicted-utility functions in sequential decision-making situations involving incidental affective states (e.g., anger, fear, economic fear) and task-related confidence. I constructed a new model combining measures to evaluate risk preferences: behavioral choices, self-reported experience self-reported experience, self-reported predicted utility, self-reported confidence. The analysis results show how domain uncertainty, framing, and emotion state of decision makers influence their subjective experience and discriminability, affective prediction, optimal decisions and exploratory regulation. I found empirically that there were significant interaction effects of framing and emotion on risk preferences: negative emotions made people more risk-averse in face of gains. When it comes to losses, anger made people more risk-averse and fear more risk seeking. I also characterized how gender and emotion influence confidence and exploratory choice behavior. The theoretical analysis nicely supports empirical findings from human experiments. The new model provides a theory that better explain and simulate human behavior under uncertainty, frames and emotions.

Thesis Supervisor: Rosalind W. Picard
Title: Professor of Media Arts and Sciences, Program in Media Arts and Sciences, MIT

Modeling and Analysis of Affective Influences on Human Experience, Prediction, Decision Making, and Behavior

by

Hyung-il Ahn

The following people served as readers for this thesis:

Thesis Reader_____
Jennifer Lerner
Professor of Public Policy and Management
Harvard University

Thesis Reader_____
Andrew Barto
Professor of Computer Science
University of Massachusetts Amherst

Thesis Reader_____
Kelly Hewett
Senior Vice President
Customer Experience Marketing Research, Bank of America

Acknowledgments

There have been a number of people without whom this dissertation could not have been completed. First of all, I wish to record my deepest gratitude to my amazing advisor, Prof. Rosalind Picard for giving me the opportunity to research emotions and decision making in the affective computing group at the Media Lab. She always imbued me with great confidence in my research. I was so lucky to get invaluable support, encouragement and guidance from her. She provided me a role model of what a great professor should be like.

I am sincerely grateful to my dissertation comittee members: Dr. Jennifer Lerner for inspiring me to do research on the role of emotions in human decision making and giving me numerous constructive inputs; Dr. Andy Barto for his insightful feedback and enriching my understanding of the reinforcement learning literature; Dr. Kelly Hewett for great supports and comments in designing and conducting human decision experiments. With my committee members' helps and supports, I was able to be inspired by diverse areas of emotions and decision making, such as psychology, behavioral economics, consumer research, machine learning and affective computing. I feel greatly honored to have this wonderful committee.

I also would like to express my deep thankfulness to Dr. Dan Ariely and Dr. Nathaniel Daw for serving as my general exam committee members. Dan's research in Predictably Irrational and Nathaniel's work in computational models of human exploratory decisions really helped me to develp a deep interest in building a computational model on subjective experience and affective prediction in human decision making under uncertainty.

I am deeply indebted to all current members and alumni in the affective computing group: Rob Morris, Ming-Zher Poh, Shaundra Daily, Rana el Kaliouby, Ehsan Hoque, Elliot Hedman, Micah Eckhardt, Daniel McDuff, Rich Fletcher, Jackie Lee, Yuta Kuboyama, Mish Madsen, Matthew Goodwin, Daniel Bender, Alex Khitrik, Kyunghee Kim, Seth Raphael, Alea Teeters, Carson Reynolds, Win Burleson, Hoda Eydgahi, Nikolaos Mavridis, Pei-Yin Chao, Ashish Kapoor, and Yuan Qi. Their advice and good friendships with them made my long academic life happier and more fruitful. In particular, I really thank Ming for providing his EDA sensors to me for my human decision experiments, Daniel (McDuff) for helping my implementation of the computerized decision tasks. Ashish and Shaundra

I dedicate this dissertation to my lovely wife, Yunjung. With her inestimable support and endurance, I was able to complete my dissertation. *With love to you.*

Contents

14

List of Figures

16

17

18

20

List of Tables

24

Chapter 1

Introduction

Human affect and emotional experience play a significant, and useful, role in human decision making and learning. Recent research in decision making and learning has integrated ideas from psychology, economics, neuroscience, and artificial intelligence and machine learning, and created new interdisciplinary areas such as behavioral economics, neuroeconomics and computational neuroscience. One of the core themes and key problems in these new areas is to verify and model the important role of subjective experience and affective prediction in decision making and learning [7, 12, 5, 41, 42, 24, 29, 36, 50].

Yet, the research for exploring and exploiting those affective and subjective influences in computational decision-making and learning models and their applications currently lags behind that in other fields. Most computational models for analyzing and simulating human experience, prediction and decision are still based on cognitive models. Traditional models based on expected utility theory assume that humans make decisions on the basis of a deliberative cost and benefit analysis. Recent models based on behavioral decision theory focuses on cognitive errors and heuristics in human judgments and decision making, but still ignore the role of emotion in human decision making [36].

This cognitive perspective is not suitable for analyzing human decision behavior in such situations where people's affective state experienced at the time of making a decision often influences their experience and prediction. For example, people have different risk attitudes and action tendencies when they are in different mood states. Incidental fear or anger influences risk perceptions and judgments [34, 33]. Investors' incidental emotion

state influences financial investment choices [36]. Therefore, for daily-life situations in which there are interactions or conflicts between affect and cognition, an affect-integrated model will be more desirable in describing human behavior than traditional cognitive models. Furthermore, the cognitive view cannot shed light on the beneficial aspects of affective decision making in synthesizing human-like decision behavior for agents interacting with humans.

The main objective of this dissertation is to redress the old cognitive view on human decision making and learning by proposing a new model, called the "affective-cognitive (AC) decision model" drawing on a growing literature offering social-psychological aspects and neural underpinnings of human decision-making behavior. The AC model assumes that an agent's decision preference for a choice mainly arises from the feedback of past subjective experiences in similar choice situations (the "experience-based" mode) and the affective prediction about future hedonic impact of choice outcomes (the "prediction-based" mode).

Human decision experiments were designed and conducted to empirically infer how people adjust the parameters (i.e., risk attitude and reference point) of their experienced-utility and predicted-utility functions in sequential decision-making situations involving incidental affective states (e.g., anger, fear, economic fear) and task-related confidence. The experimental analysis and results show how domain uncertainty, frames, and emotion state of decision makers influence their subjective experience and discriminability (i.e., the level of easiness in discriminating which choice has a greater overall outcome), affective prediction, optimal decisions and exploratory regulation (i.e., trade-offs between exploration and exploitation). The theoretical analysis nicely supports empirical findings from human experiments. Figure 1-1 represents the main topics of the research, that is, decision making under uncertainty, frames and emotions.

1.1 Affective-Cognitive (AC) Decision Model

The AC model extends and integrates existing theories of prospect theory (PT) [29, 61, 30] and reinforcement learning (RL) [58], drawing on a growing literature offering the role of affect in decision making and the neural underpinnings of human decision behavior. The pro-

28

Decision Making under Uncertainty, Frames, and Emotions
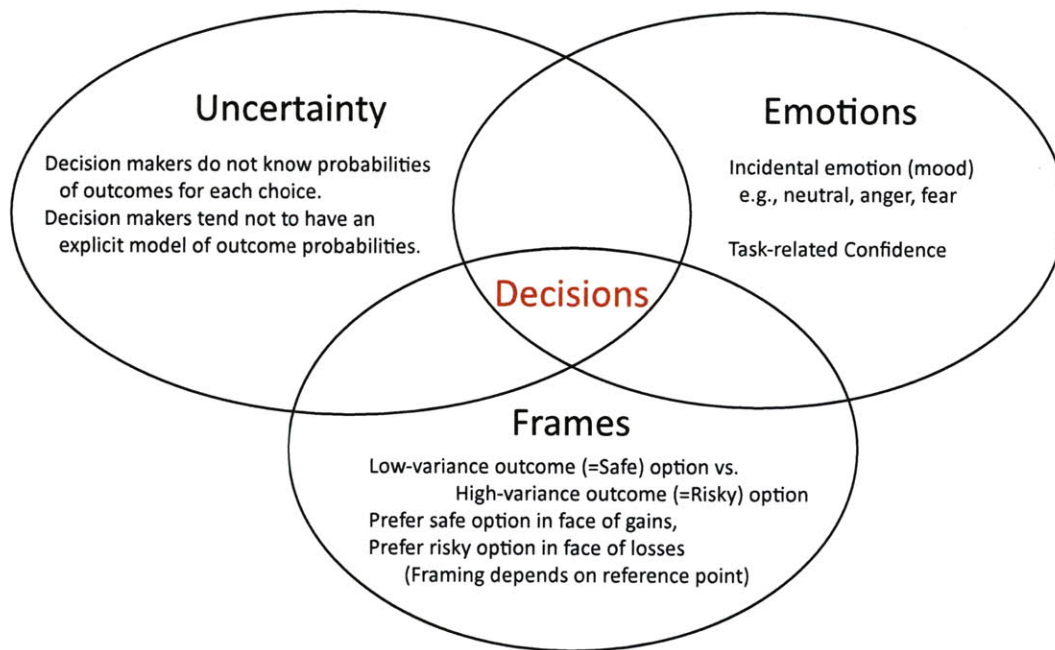
Figure 1-1: The main topics of the research: decision making under uncertainty, frames and emotions

posed framework assumes that an agent's decision preference for a choice mainly arises from the feedback of past subjective experiences in similar choice situations (the "experience-based" mode) and the affective prediction about future hedonic impact of choice outcomes (the "prediction-based" mode). Broadly speaking, the experience-based mode is a reflexive, associative and automatic process, whereas the prediction-based mode is a goal-directed reflective process [1, 14, 29, 50]. The AC model elucidates how the current state–including both cognitive and affective states–systematically influences the computation of decision utility of a choice through the experience-based and prediction-based modes. The AC model employs Prospect Theory (PT)-based parameterized subjective value functions to model people's experienced-utility and predicted-utility functions. It assumes that the shapes (or parameters) of these subjective value functions dynamically vary with the decision-maker's task-related and/or incidental affective states in sequential decision making.

The PT-based parameterized subjective value functions (i.e., experience-utility function,

predicted-utility function) in the AC model can realize different risk attitudes and framing varying with task-related affective states and how this setup helps regulate the trade-off between exploration (seeking new information) and exploitation (seeking to maximize overall utility under current information) in uncertain domains and offer great overall performance for fewer resources. For example, in one of the states I model, a positive confidence state (i.e., positive goal-achieving state, being very confident in the current task), the learner becomes more sensitive to both likely gains and likely losses. The result facilitates computational exploitation, and can be expressed in terms of the temperature in a Boltzmann model as affect-parameterized cooling. Similarly, there is a negative confidence state (i.e., negative goal-achieving state, being not at all confident in the current task), where an agent becomes less sensitive to both likely gains and likely losses, which facilitates exploration and appears as affect-parameterized heating. Thus, instead of temperature being a scalar value following some kind of (typically fixed, monotonic) annealing schedule, the affective state can drive it up or down, regulating it based on moment-by-moment circumstances.

The AC model describes how people's subjective experience and reference point selection (framing) influence their "subjective discriminability" (i.e., the level of easiness in discriminating which choice has a greater overall outcome), exploration-exploitation regulation and optimal decisions. Computational simulations demonstrate that subjective discriminability can be increased by the use of a PT-like experienced-utility function, the shape of which reflects people's loss-averse attitude, risk-averse attitude when it comes to their gains, and risk-seeking attitude when it comes to their losses, with an appropriate reference point. An increased subjective discriminability enables an affective-cognitive agent employing the nonlinear experienced-utility function to need fewer exploratory trials for optimal decisions, compared to the rational agent employing the linear experienced-utility function. That is, an affective-cognitive agent with a good framing can achieve greater subjective discriminability and quicker optimal decisions under uncertainty than a rational agent whose discriminability (called "objective discriminability") is always the same regardless of the reference point selection (gain or loss framing).

The incidental emotion state (e.g., a mood state) can influence the risk attitude (risk seeking or risk averse) adjusting the parameter values (the shape and the reference point)

of the subjective value functions. Different affective states such as anger, fear, sadness and hope may lead to different subjective discriminabilities for the same underlying objective outcome distribution.

I present two-armed bandit-task experiments for humans as a framework to infer how people's incidental emotion influences their subjective experience (experienced utility function), confidence and affective prediction (predicted utility function) in decision making under domain uncertainty, framing and emotions, based on people's actual choice behavior and self-reported data. Also, the AC model was used to infer subjective value functions explaining human behavioral and self-reported data. Human decision experiments confirmed the validity of simulation results of the AC model about framing and emotional effects on subjective discriminability and optimal decisions, i.e., the interactions of framing and emotion in uncertain domains influenced people's subjective discriminability, predictions and optimal decisions.

1.2   Dissertation Contributions

This dissertation presents a new computational perspective on the role of subjective experience and affective prediction in human decision making and learning, drawing on the findings in diverse areas of decision science as behavioral economics, neuroeconomics, psychology and machine learning.

The proposed framework of human behavioral experiment and computational analysis shows how we can observe and infer the subjective and affective influence on decision making under uncertainty, frames and emotions. This research empirically infers how people adjust the parameters (i.e., risk attitude and reference point) of their experienced-utility and predicted-utility functions in sequential decision-making situations. Also, I present how we can computationally model subjective experience and affective prediction to build artificial agents with quicker optimal decisions in learning and decision making under uncertainty. The AC model will be able to regulate the agents' risk attitude (risk seeking or risk averse) and reference point selection with a simulated affective state in order to achieve increased adaptability and robustness. In more detail, the dissertation

31

- Defined subjective discriminability. Showed both computationally and empirically that bigger subjective discriminability leads to more optimal decisions.

- Characterized how subjective and affective influences may help or harm human decision making depending on domains, frames, emotions and their interactions.

- Constructed a new model combining measures to evaluate risk preferences: behavioral choices, self-reported experience (subjective discriminability), self-reported predicted utility (predicted-utility difference), self-reported confidence.

- Introduced two different kinds of subjective value functions (experienced-utility (EU) function and predicted-utility (PU) function) whose parameters change with emotions and provided a method to infer PU and EU functions from self-reported EU and PU data in each emotion condition.

- Showed how to compute reference points (EU frame, PU frame) for each of the EU and PU functions.

- Analyzed risk attitudes based on EU and PU frames as well as on the frame given by experimenter.

- Observed how emotions influence the reference point selection (framing).

- Discovered the frame and emotion effects (main and interaction effects) in decision making under uncertainty.

- Measured how experience, gender and emotion influence confidence and prediction.

- Introduced the confidence-dependent predicted utility function.

- Presented a new emotion-refresher method.

- Defined the value of risk (VOR).

- Characterized how domain, frame, emotion influence decision making: Negative emotions in face of gains (more risk-averse), Anger in face of losses (more risk-averse), Fear and Economic fear in face of losses (more risk-seeking).

- Showed how human behavior can be described by emotion-shaped EU and PU functions.

- Provided a theory that better explain/simulate human behavior under uncertainty, frames and emotions.

1.3 Dissertation Outline

The contents of the remaining chapters are as follows:

Chapter 2. Background and Proposed Model covers some relevant concepts and approaches in prospect theory (PT), reinforcement learning (RL) and psychological theories on the role of subjective experience and affective prediction in decision making. Then it presents the essential features of the affective-cognitive (AC) decision model, and compares and compares the AC model with other existing models.

Chapter 3. Modeling Subjective Experience and Affective Prediction describes the computational model of subjective experience and affective prediction in the AC model employing the PT-based parameterized experienced-utility and predicted-utility functions, and defines a new concept called "subjective discriminability". Compares this with the objective discriminability. Computational simulations confirm that framing (reference point selection) influences subjective discriminability, and that subjective decision agents (relying on subjective experiences) can achieve quicker optimal decisions in simple learning situations such as two-armed bandit tasks, compared to objective decision agents (relying on objective outcomes).

Chapter 4. Human Decision Experiments: Method and Hypotheses describes the experimental design and procedure of human decision experiments (based on two-armed bandit tasks) conducted to observe and infer the subjective and affective influence on decision making under uncertainty, frames and emotions. It also presents the main hypotheses behind experiments.

Chapter 5. Human Decision Experiments: Analysis and Results details the analysis and results of human decision experiments. The analysis confirms main hypotheses of the experiments.

Chapter 6. Discussion and Future Work summarizes the dissertation and suggests future work.

Chapter 2

Background and Proposed Model

The proposed affective-cognitive (AC) decision model extends and integrates existing theories of prospect theory (PT) and reinforcement learning (RL), drawing on a growing literature offering the role of affect in decision making and the neural underpinnings of human decision behavior. Prospect theory (PT) mostly focuses on people's subjective prediction and choice behavior in one-shot decision situations under risk (i.e., situations in which people make a one-shot choice with full knowledge of the outcome probability distributions of options). Although some behavioral decision-making models [20, 22, 23] and reinforcement learning models in computational learning theory [16, 56, 58] have been applied to analyze people's behavior in decision-making situations under uncertainty, they have not fully incorporated the characteristics of PT-based subjective value functions (i.e., risk attitudes depending on reference-point dependency, diminishing sensitivity, loss-aversion) for modeling people's subjective experience and affective prediction. The AC model uses PT-based parameterized subjective value functions to model people's experienced-utility and predicted-utility functions. It hypothesizes that the shapes (or parameters) of these subjective value functions adjust to a decision-maker's task-related and/or incidental affective states on the fly during a sequence of decision-making trials.

The AC model adopts Kahneman's utility taxonomy [28] and combines the different concepts of utilities to provide a computational decision-making framework for more general situations. Kahneman's utility taxonomy is useful for distinguishing multiple concepts of "utility." In modern economics, the utility of outcomes usually refers to their weight

in decisions: utility is inferred from observed choices and in turn used to explain choices. This behavioral concept of utility is called decision utility. Kahneman distinguished experienced utility from decision utility, although he did not propose a computational model that combines the two kinds of utility concepts. Experienced utility refers to the experiences of pleasure and pain, as Bentham used it [31]. It is the affective or hedonic impact of an obtained outcome after a choice. Also, predicted utility is a belief about the future experienced utility of a choice before making a decision. Also, recent findings in neuroscience suggest that the neural substrates of liking (pleasure) are separate from those of wanting (motivation) in the human brain [6], so there is evidence from neuroscience that supports treating these concepts differently when modeling human decision making. Kahneman's concepts of experienced utility and decision utility are associated with liking and wanting, respectively [7].

2.1 Prospect Theory (PT): subjective influence on predictions and decisions

Prospect theory (PT) [30] was designed to describe people's decision-making behavior under risk (i.e., one-shot choice in a situation where outcomes are "uncertain" but the outcome probability distributions of candidate options are "known" to the decision maker). For instance, in the domain of potential gains, as in the situation where one option yields a sure-outcome of winning $5 ((outcome, Pr) = ($5, 1.0)) and the other option has a 50% chance of winning $10 and 50% chance of getting nothing ((outcome, Pr) = ($10, 0.5; $0, 0.5)), people tend to prefer the sure-outcome situation to the gamble or variable-outcome situation of equal expected value. People don't want to bet on an uncertain choice that might spoil their chance at a gain. That is, people are "risk-averse in the domain of likely gains." In the domain of potential losses, however, as in the situation where one option yields a sure-outcome of losing $5 ((outcome, Pr) = (-$5, 1.0)) and the other option has a 50% chance of losing $10 and 50% chance of getting nothing ((outcome, Pr) = (-$10, 0.5; $0, 0.5)), people prefer the gamble of equal expected value to the sure-outcome situation. People want to bet on the choice that might mitigate the loss. That is, people are "risk-

when it comes to losses.

uncertainty   frames   emotions

subjective experience   affective prediction

decision

Reinforcement Learning (RL) with a linear utility function	influence of past experience on decisions (multiple trials in decision under uncertainty)	*v* Objective value (Utility) / *x* Outcome
Prospect Theory (PT)	influence of frames and risks on decisions (one-shot choice in decision under fully known outcome distributions)	*v* Subjective value (Utility) / Losses Gains / $x - x_{ref}$ Relative outcome
Affective-Cognitive (AC) model	emotion state-based parameterized experienced-utility (EU) and predicted-utility (PU) functions to model the role of emotions and frames in decision under uncertainty	*v* Subjective value / Losses Gains / $x - x_{ref}$ Relative outcome / EU and PU functions have emotion state-dependent shapes. Also, they are reference-point dependent.

Figure 2-1: Comparison in terms of utility functions: RL with a linear utility function, PT, AC model

seeking in the domain of likely losses." Also, people are "more sensitive to potential losses than potential gains" or show "loss-averse" attitudes. For instance, most people would not take a gamble with a 50% chance of winning \$10 and 50% chance of losing \$10. Loss

v · Subjective value (Utility)

Losses | Gains

$x - x_{ref}$
Relative outcome

$$v(x - x_{ref}) = \begin{cases} (x - x_{ref})^{a}, & x - x_{ref} \geq 0 \\ -\lambda(-(x - x_{ref}))^{b}, & x - x_{ref} < 0 \end{cases}$$

$$0 < a < 1, \quad 0 < b < 1, \quad \lambda > 1$$
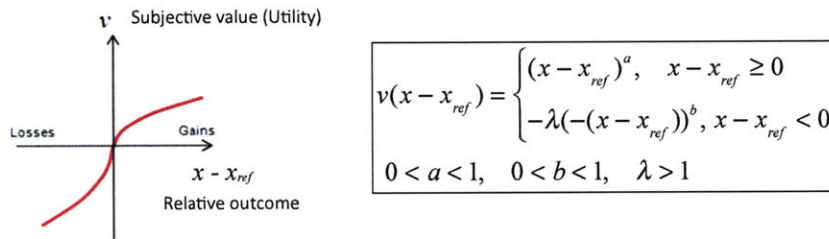
- Reference Dependence:
gains and losses are defined relative
to the reference point (x_{ref})

- Diminishing sensitivity:
decreasing marginal utility
Concave above the reference point
Convex below the reference point

- Loss aversion:
the function is steeper in the loss frame

Figure 2-2: The Prospect Theory (PT) Value Function, Kahneman and Tversky (1979)

aversion refers to the tendency for people to strongly prefer avoiding losses to acquiring gains.

PT describes human behavior with a subjective value function. This function has three essential characteristics. First, gains and losses are defined relative to a reference point (reference dependence). If an outcome x is greater or smaller than a reference point x_{ref}, the relative outcome $x - x_{ref}$ is viewed as a gain or a loss, respectively. The reference point might depend on the decision maker's expectation of the average outcome over relevant choices. Second, the function has a concave form in the domain of gains and a convex form in the domain of losses (diminishing sensitivity). Third, the function is steeper in the negative domain than in the positive domain (loss aversion). These three characteristics describe how people make a prediction on the hedonic impact ("experienced utility") of the future outcome of a candidate option. This cognitive belief or expectation on the future experienced utility is called "predicted utility." That is, the PT subjective value function was originally constructed as a predicted-utility function, which describes the decision maker's predictions before a decision (i.e., in the predicting phase). Note that "predicted utility" is distinguished from "experienced utility" which means the actual hedonic impact of the outcome obtained after a decision (i.e., in the experiencing phase).

Research in PT, however, mostly focuses on the single-stage unambiguous decision-making problems where future outcomes are uncertain but the outcome probability distributions of candidate actions are known to the decision maker, rather than on sequential decision-making problems involving learning in unknown stochastic domains such as in the Reinforcement Learning (RL) literature. In other words, existing work in prospect theory has focused mostly on non-sequential decision scenarios with known outcome distributions.

While PT has been developed in economics for domains with known outcome distributions, the AC model enables PT to be used in decision making and learning for unknown and changing stochastic outcome distributions. The AC model will thus extend PT in that it explains the human decisions PT explains and also learning behaviors in more complex and realistic situations. Also, while PT just describes human decision behaviors, the AC model can both describe and synthesize behaviors.

2.2   Reinforcement Learning (RL): decision making under uncertainty

Most traditional work on RL and Markov decision processes (MDP) has assumed that agents make decisions to maximize the expected total reward and that all the reward signals influencing an agent's behavior come from the environment and depend on only the external state. Yet, "intrinsically motivated RL framework" assumes that an agent's reward signals are determined by processes within its brain that monitor not only external events through exteroceptive systems but also the agent's internal state, which includes information pertaining to critical system variables as well as memories and accumulated knowledge [3, 55, 56, 57]. The AC model can be viewed as consistent with the intrinsically motivated RL framework. The new model, however, focuses on how the form of the utility function as the critic changes with the agent's internal state (e.g., affective state) and influences decisions and learning. Note that most traditional RL simulations often involve a linear utility function (reward function) to evaluate outcomes or just use a fixed form of the utility function without focusing on how the changes in the shape of utility function (i.e., changes in sensitivities to the expected gains and gains) have an impact on the agent's

learning and decision making.

A few RL and MDP models have taken into account both the expected total reward and the risk by transforming the total reward by exponential utility functions (Howard and Matheson, 1972; Koenig and Simmons, 1994), by applying the worst-case criterion for total discounted return (Heger, 1994), or by transforming temporal differences through a risk-sensitive linear utility function (Mihatsch and Neuneier, 2002). Yet, none of the existing methods are informed by recent Nobel-prize winning findings in behavioral economics–namely prospect theory (PT), which addresses important subjective effects known to influence human decision making. Moreover, none of the existing approaches discuss how the current affective state at the time of decision making can influence decision making and help regulate the trade-off between exploration (i.e., selecting a currently-estimated suboptimal action to gather new information on that action) and exploitation (i.e., selecting a currently-estimated best action).

A variety of reinforcement learning (RL) algorithms for sequential decision making under uncertainty have been used to model learning and decision making in situations where the agent initially does not know the exact outcome distributions of candidate options and should learn them by trials (these sorts of learning problems are often called "decision making under ambiguity (Bechara and Damasio, 2005; Glimcher and Rustichini, 2004)" or "decision making from feedback (Barron and Erev, 2003)" in decision-making literatures). There are two distinct kinds of RL approaches: model-free RL and model-based RL [14, 58]. The model-free RL, such as Q-learning algorithms, updates a cached value (or a weighted average outcome) for each state-and-action pair, and then, whenever a new experience happens in a similar situation, it updates the cached value by a new outcome. It does not keep any models of state transitions and outcome probability distributions. Thus, the model-free RL is computationally simple but comes at the cost of inflexibility. In contrast, the model-based RL estimates the explicit models of state transitions and outcome probability distributions. Thus, it allows the agent to make predictions on outcome values of alternatives on the fly, and to quickly adapt to the dynamic change of the environment. The model-free and model-based RL approaches, respectively, can be associated with the models of the automatic affective habit-based control and reflective goal-directed control [14]. In

40

the perspective of RL algorithms, the experience-based mode can be linked to the model-free RL and the prediction-based mode can be linked to the model-based RL.

Although RL is generally treated as a computational theory within the machine learning and operation research communities, there have also been some attempts to apply RL principles to analyze human decision behavior [9, 15, 19]. Note that the existing RL theories have not been developed to describe human behavior. Therefore, the RL theories have not examined the subjective and affective influences on human prediction and experience of choice outcomes. In contrast, the AC model draws on subjective experiences and affective prediction. As the first step of the AC model development, I use the Prospect Theory (PT)-based parameterized subjective value functions in order to model the subjective characteristics of human decision making. I assume that the shapes (or parameters) of subjective value functions dynamically vary with a decision-maker's task-related and/or incidental affective states.

2.3 Affective influence on decisions

2.3.1 Integral emotion, incidental emotion, and task-related emotion

Expected emotion (cognition) and immediate emotion (genuine emotion): In the affective decision-making literature, predicted utility is often called expected emotion (a.k.a. anticipated emotion) [36]. That is, expected emotion means a cognitive belief on the affective impact of future outcomes (i.e., a prediction on the future experienced utility of outcomes). Note that expected emotion is a cognitive belief, not a genuine affective response. Compared to expected emotion, immediate emotion is a genuine affective response.

Different types of immediate emotion: There are three types of immediate emotion experienced at the time of decision making: integral emotion (a.k.a. anticipatory emotion), incidental emotion, and task-related emotion [36, 38, 12]. Cohen *et al.* [12] classified them as follows: "*Integral emotion* is affective responses that are genuinely experienced and directly linked to the object of decision. These integral affective responses include momentary feelings elicited by features of the object, whether these features are real, perceived, or only

imagined. *Incidental emotion* is affective responses whose source is clearly unconnected to the object to be evaluated. In addition to a person's current mood that is typically unrelated to the decision at hand, incidental emotion may also come from a person's emotional dispositions and temperament, or from any contextual stimuli associated with integral emotion. *Task-related emotion* is affective responses that are elicited by the task or process of making decisions, as opposed to direct, integral responses to features of the target objects or purely incidental feelings. For example, the emotional stress of having to choose between two very attractive offers would be considered task-induced in that it is the process of having to choose between these two offers that is stressful, not the offers themselves."

It appears that there are several mechanisms involved in how emotions influence decisions. First, *Integral emotion* provides target evaluation either through a noninferential, associative and automatic way (e.g., approach/avoidant action tendency and an embodied mode of evaluation) [41, 4, 44] or through an inferential and reflective way such as the "how-do-I-feel-about-it?" heuristic [51, 43]. Second, *Incidental emotion* influences judgments through the cuing of mood-congruent thoughts ("affective referral", noninferential, associative and automatic accesses to a prior evaluation of the target stored in memory) [25] or affect-as-information mechanisms ("affective coloring", inferential and reflective processes involving a constructive search for additional information) [10, 43, 46]. According to the mood-congruency hypothesis [25], mood states (pleasant or unpleasant) automatically cue similarly valenced materials in memory, thereby biasing people's perceptions of the target at the time of evaluation. Also, the affect-as-information hypothesis [10] is that people often examine their feelings in their evaluation, and positive or negative feelings from the mood state at the time of judgment may be misattributed to the target. Also, this misattribution tend to disappear when people recognize the true source of incidental affect. In the AC model, the automatic and associative processes are linked with the experience-based mode, and the inferential and reflective processes are with the prediction-based mode.

2.3.2 Incidental affective influence on risk attitudes

I begin with a review of some findings on the literature already.

Positive incidental emotions

When the decisions involve low risks and low stakes, positive mood individuals tend to have more optimistic (mood-congruent) expectations about the outcomes and, therefore, take greater risks compared to neutral mood individuals. Yet, when the stakes are high and the potential for losses significant, positive mood individuals become risk-averse because they want to sustain their positive affective state, which a loss would break. That is, compared to neutral moods, positive moods promote risk-taking attitude in the domain of likely gains (or in situations with only upsides) but risk-avoidance in the domain of likely losses (or in situations with only downsides) [12, 25, 26].

Negative incidental emotions

People's risk attitude under negative affective states is not only a function of the level of arousal associated with the affective state, but also a function of the appraisal content of the affective state. Thus, different emotions with the same valence but differing appraisals, such as fear and anger, influence people's risk perception in different ways [32, 33, 34, 46].

Fear and Anger: Although fear and anger are both high-arousal negative emotions, fear triggered in one situation tends to evoke more pessimistic risk estimates and risk-averse choices in other unrelated situations, whereas anger does the opposite, evoking optimistic risk estimates and risk-seeking choices [32, 33, 34]. This is because fear is typically coupled with situations of uncertainty and low control, whereas anger is typically coupled with situations of certainty and high control.

Anxiety and Sadness: Raghunathan and his colleagues [46, 47] found that in gambling decisions (a gamble offering a 6/10 chance of winning $5 vs. a gamble offering a 3/10 chance of winning $10) and in job selection decisions (an average salary with high job security vs. a high salary with low job security) involving low-risk/low-reward and high-risk/high-reward options, anxious individuals tend to prefer low-risk/low-reward options and sad individuals tend to prefer high-risk/high-reward options. The authors state that this is because anxiety, which is typically coupled with situations of low control and high uncertainty, triggers a

goal of risk and uncertainty minimization, whereas sadness, which is typically experienced in response to the loss of a source of reward, triggers a goal of reward maximization. In other words, because, even though their states are incidental, sad individuals tend to infer that they have lost something of value, which activates a goal of reward acquisition that shifts preferences toward high-reward options. In contrast, anxious individuals tend to infer that the situation is uncertain and beyond control, which triggers a goal of risk avoidance that shifts preferences toward low-risk options.

Also, Lerner *et al.* [35] found that incidental states of sadness reverse the classic endowment effect, that is, the tendency to place a higher value on objects that are already in our possession compared to identical objects that not in our possession. According to the authors, this is because sadness creates a motivation to change the current situation, which increases the willingness to pay for objects that are not in our possession (higher purchase prices) and also increases the willingness to sell objects that currently are (lower selling prices).

Disgust: According to Lerner *et al.* [35], incidental states of disgust eliminate the endowment effect. That is, disgust triggers an impulse to get rid of objects that are currently in our possession (lower selling prices) without necessarily distorting the value of objects that are not in our possession (unchanged purchase prices).

2.4 The role of subjective experience in decision making under uncertainty

The human adaptive learning studies on a task known as the Iowa gambling task (IGT) have shown that people with a normal decision-making ability are good at quickly detecting the optimal decision in the task, whereas patients with emotional deficits related to the vmPFC damage are not [4, 5, 40]. The IGT involves repetitive trials in each of which the participant selects a card among four decks of cards and obtains a monetary outcome written in the selected card. Participants have no initial information about the underlying outcome distributions of four decks. Decision making under uncertainty in this sort of

44

situation is often called "decision-making under ambiguity," distinguished from decision making under the full information of underlying outcome distributions. Two decks produce higher regular value but occasionally with big penalty so they are risky and disadvantageous (or lower expected outcome), whereas the other two decks produce lower regular value but occasionally with lesser penalty so they are relatively non-risky and advantageous (or higher expected outcome). Compared with patients with prefrontal damage, after experiencing a big penalty, normals began to generate anticipatory skin conductance arousal whenever they made a choice from the decks that would turn out to be risky, even before they explicitly recognized that it was risky. Yet, vmPFC patients never developed anticipatory arousal although some eventually realized which choices were risky. Overall, vmPFC patients were found to select more from risky decks, obtaining less overall outcome than normals.

Glimcher and Rustichini [21] have pointed out the following: First, "patients with vmPFC lesions seem to lack an aversion to ambiguity or losses that normal participants have, an aversion that may be quite advantageous under many conditions." Second, "the process of learning and evaluating feedback may involve emotion-related areas." Third, "the process of decision making under ambiguity may be very different from when participants simply choose between options without any feedback or learning taking place at the same time."

Also, Pham [44] has suggested the following: First, "in tasks that do not involve outcome feedback, VMPC patients and normal participants exhibit comparable levels of risk-seeking and impulsivity. This suggests that presumably emotionally impaired VMPC patients are not inherently more risk-seeking and impulsive; rather they differ in how they respond to and learn from outcome feedback." Second, "it is well established that integral affective responses to a target that are positive generally trigger approach tendencies, whereas those that are negative generally trigger avoidant tendencies, even if descriptions of the targets and their cognitive assessments are held constant." Third, "integral affective responses often serve as distinct proxies for value. What the Damasio studies [5], along with other studies, suggest is simply that integral-affect-motivated approach and avoidance – that is, affective behavioral regulation – is very sensitive to emotion-producing outcome feedback."

Consistent with these researchers' hypotheses, the AC model has a view that antic-

ipatory arousal relates to ambiguity aversion and is an indicator of the influence of the experience-based mode involving experience-based emotional feedback and integral (anticipatory) affective responses on human decision making.

2.5 Proposed Model: Affective-Cognitive Decision Model

The Affective-Cognitive (AC) model proposed in this dissertation extends Prospect Theory (PT)-based subjective value functions to model people's experienced-utility and predicted-utility functions. It is assumed that the parameters of these subjective value functions can change with the current affective state in order to include the subjective and affective influences such as loss aversion, the reference point dependency (framing effect) and risk attitudes in adaptive decision making. Experienced utility (liking), which is associated with the consummatory and hedonic properties (pleasure or displeasure) of current outcomes of a choice, arises from the evaluation of our experience of the outcomes. Thus, this is about how much people like or dislike the outcomes of our choice. Decision utility (wanting), which is associated with the appetitive and motivational properties of future expected outcomes of a choice, relates to the degree to which a choice is selected, and can be associated with and inferred from actual observed behaviors. Also, it is very important to note that decision utility (wanting) does not result from a unitary process [7, 28].

The experience-based and prediction-based modes of the AC model can be linked to the model-free RL and model-based RL approaches [14, 58], respectively. In contrast to the traditional RL approaches, however, the AC model incorporates subjective experiences and affective prediction by means of subjective value functions. The experience-based mode employs the experienced-utility function to evaluate the subjective (or hedonic) impact of resulted outcomes, and the predicted-based mode capitalizes on the predicted-utilty function to allow for the affective influence on the prediction about future outcomes.

The AC model aims to model the multiple pathways of how the current state–including both affective and cognitive states–systematically influences decision utility. The AC model assumes that the agent is in a certain cognitive and affective state at the time of decision making, and that the decision utility for a candidate option in the current state is composed

	Experience-based learning under uncertainty	Experienced utility (EU) function	Reference point dependency (framing)	Artificial agent simulation	Predicted utility (PU) function	Emotional influence on utility functions & reference points
Prospect Theory (PT)	No one-shot decision under known risks	No	Yes modeled for the PU function, a fixed reference point	No	Yes assumed to be a fixed shape	Yes modeled for the PU function, a fixed reference point
Reinforcement Learning (RL)	Yes multiple trials under unknown and uncertain outcomes	Yes often assumed to be linear, no reference point modeled	No	Yes objective discriminability based	No	No
Affect-Cognitive (AC) model	Yes multiple trials under unknown and uncertain outcomes	Yes PT-like nonlinear subjective value function, reference point dependent	Yes modeled for PU and EU functions each, reference point may change during the task	Yes subjective discriminability based	Yes the shape changing with the decision maker's confidence	Yes modeled for PU and EU functions each, reference point may change during the task

Figure 2-3: Comparisons of PT, RL, and AC models

of two main influences: the experience-based mode and the prediction-based mode.

Figure 2-3 shows comparisons of PT, RL and the proposed AC model. The following sections will discuss main features of the AC model such as:

- The experience-based mode and total-experienced utility

- The prediction-based mode and predicted utility

- The experienced-utility function differs from the predicted-utility function

- Decision utility and the tradeoffs between affect and cognition

- Affective shaping can model incentive salience

- Confidence state (goal-achieving state) influences prediction

- Incidental emotion state and framing influences experience, prediction and risk attitude

2.5.1 The experience-based mode and total-experienced utility

Past affective experiences associated with a candidate option in similar situations to the current state are automatically retrieved from the episodic memory and reactivated in the short-term memory [5, 41]. This overall reactivation contributes to the motivation of selecting the option. In the AC model, the "experience-based mode" refers to the overall valence-based automatic affective influences.

The experience-based mode in the AC model can be approximated by model-free caching reinforcement learning (RL) algorithms [58]. This computation method is also similar to Kahneman's moment-based approach to evaluate a statistically aggregated overall value (e.g., the recency-weighted average value) over past experienced utilities (past moment likings). According to Kahneman's definitions on different utilities [28], this statistically computed value is called "total-experienced utility" (a.k.a. "total utility"). Note that the concept of total-experienced utility relates to the overall liking of past experiences, whereas that of experienced utility relates to the moment liking of an experienced outcome. Total-experienced utility or the experience-based mode in the AC model explains the role of past experiences in the computation of decision utility (wanting). It can be associated with "action value" in model-free RL and "anticipatory emotion" in the affective decision making literature [4, 12, 36, 44].

The experience-based mode keeps a cached value for each state and option pair, and then, whenever a new experience happens in a similar situation, it updates the cached value (total-experienced utility) by a new experienced utility (moment liking). In other words, this mode computes the weighted average over the past experienced utilities. The "experienced-utility function" is employed to compute the experienced utility (moment liking) for an experienced outcome. This function is modeled by a Prospect Theory (PT)-based parameterized subjective value function whose shape changes with the experiencing affective state, that is, the affective state at the moment of actual consummatory experience.

2.5.2 The prediction-based mode and predicted utility

The prediction about the affective experience of future outcomes of a candidate option contributes to the motivation of selecting the option. In the AC model, the influence of prediction on the motivation of behaviors is called the "prediction-based" mode. Note that the cognitive belief on future experienced utility is influenced by the current affective state (i.e., drive states, anticipatory, incidental and task-related affective states, etc.) as well as the current cognitive state (i.e., goal-relevant environmental stimuli and cognitive goals). Thus, this prediction-based mode includes both deliberative and affective processes. The prediction-based mode can be associated with Kahneman's definition on "predicted utility" [28].

In the AC model, computing the predicted utility for a candidate option is approximated by two phases: the "deliberative estimation" phase and the "affective shaping" phase. Although the estimation of the future long-run outcome distribution for a choice is very deliberative and goal-directed (i.e., deliberative estimation), the ultimate prediction is influenced by the current affective state at the moment of prediction (i.e., affective shaping).

Deliberative estimation attempts to estimate the long-run objective outcome distribution (costs and benefits in terms of goals) for a candidate option. From a theoretical perspective, any model-based learning method can be employed if it is useful for approximating the distribution. In contrast to traditional RL algorithms in which the expected future outcome is directly used for the action-selection model (decision-making policy), the AC model computes predicted utility, which incorporates the affective influence on the prediction. This additional step, that is, transforming the long-run outcome distribution into a predicted utility, is called the "affective shaping" phase.

Affective shaping approximates the influence of the current predicting affective state on the predicted utility. This step takes cognitive beliefs on the long-run outcome as an input, and transforms those expected outcomes into predicted-utility samples through the predicted-utility function. The predicted utility is the average of the predicted-utility samples.

In the AC model, the predicted-utility function is another PT-based parameterized subjective value function, which is distinguished from the experienced-utility function.

49
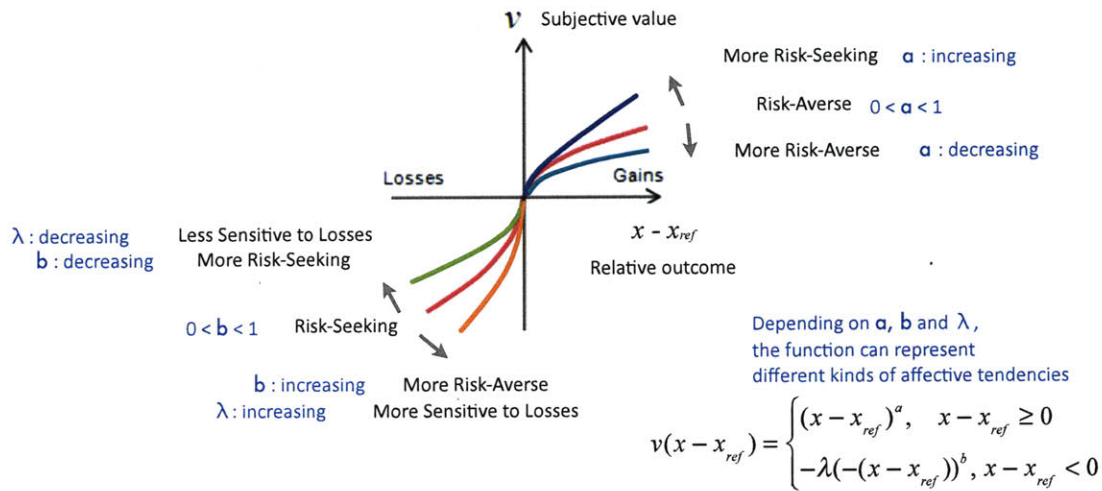
Figure 2-4: Risk attitudes (risk seeking or risk averse in the domain of likely gains and losses, sensitivity to losses) depend on the values of the PT parameters

Figure 2-4 shows how the PT value function (predicted utility function) is parameterized to describe different risk attitudes including the risk attitudes (risk seeking or risk averse) in face of gains and losses, and the sensitivities to losses.

2.5.3 The experienced-utility function differs from the predicted-utility function

Note that AC model employs two different kinds of subjective value functions for a certain kind of outcome, an "experienced-utility function" and a "predicted-utility function," although both functions are modeled by PT-based parameterized functions. The experienced-utility function (used in the consummatory, experience stage) is to evaluate moment-experienced utility for an obtained outcome, whereas the predicted-utility function (used in the preparatory, prediction stage) is to modulate a deliberative belief on the future outcome into an affective value or predicted utility in order to compute the expected hedonic impact of the future outcome. The AC model assumes that the subjective value function employed in evaluating current experience (i.e., experienced-utility function) is different from that employed in predicting future experience (i.e., predicted-utility function).

In the affective decision making literature "projection bias" tells that people often

project their current affective state (i.e., the affective state at the time of prediction (decision making)) into their prediction because they fail to correctly predict their future experiencing affective state (i.e., the affective state at the time of future experience) [24]. In particular, when people have not had much previous experience in similar decision-making situations, they often make their prediction as if their future experiencing affective state were the same as their current affective state at the time of prediction. Yet, it is very often that the affective state at the time of actual experience in the future (employed in the experienced-utility function) is different from the affective state at the time of prediction (employed in predicted-utility function). Thus, this supports the assumption of the AC model that there are two separate subjective value functions for experience and prediction.

2.5.4 Decision utility and the tradeoffs between affect and cognition

The AC model assumes that decision utility for a candidate option is composed of total-experienced utility (experience-based mode) and predicted utility (prediction-based mode). Psychological experiments have shown that cognitive load at the time of decision making as well as personal disposition may influence the trade-off between the influences of the experience-based mode and the prediction-based mode and the trade-off between the goal-directed deliberative estimation and the affective shaping [53]. In other words, when a person is stressed or more cognitively loaded, he or she tends to rely more on the experience-based mode and may be more strongly influenced by the affective shaping.

Lowenstein and O'Donoghue's decision-making model [37] assumes that one mode totally relates to the affective system and the other mode totally relates to the deliberative system. In the AC model, however, both experience-based and prediction-based modes are influenced by the current affective state in its own way, and the prediction-based mode includes deliberative estimation. The experience-based mode has direct influence through the caching mechanism where total-experienced utilities for distinct affective states are stored independently, whereas the prediction-based mode has indirect influence through the affective shaping.

2.5.5 Affective shaping can model incentive salience

In the computational-modeling perspective, incentive salience (cue-triggered 'wanting') can be also modeled by affective shaping, which uses a parameterized predicted-utility function for each kind of outcome. Incentive salience is irrational motivation, distinguished from the goal-directed prediction. Berridge *et al.* [6, 7] says, "Attribution of incentive salience transforms mere sensory information about rewards and their cues (sights, sounds and smells) into attractive, desired, riveting incentives. Its attribution to a percept or other representation is to make it a 'wanted' target of motivation. For instance, if drug cues trigger activation of sensitized mesolimbic dopamine systems, then an addict may be moved to take drugs again by hyper-incentive wanting. In general, the incentive salience hypothesis and relevant experimental results are fully supportive of Loewenstein's visceral factors theory, and provide one specific mechanism by which visceral factors might actually overwhelm volition to produce irrational choices." The computational method of incentive salience in the AC model involves the same "affective shaping" phase as in that of predicted utility. Yet, the incentive salience mechanism (where "irrational" wanting is triggered by activation of the brain dopamine system) should be viewed as an independent pathway into decision utility. This is consistent with Berridge's view. For instance, experienced people (e.g., experienced addicts, compared to first-time drug users) usually know that they will not actually 'like' the outcomes related to cues later, but they still 'want' and work for the outcomes.

2.5.6 Confidence state (goal-achieving state) influences prediction

The AC model implements one specific kind of task-related affective state, called the "confidence state (goal-achieving state)". The model assumes that the confidence state is associated with how much experienced utility (pleasure or displeasure) the agent has achieved in terms of its goal, and that the confidence state is a measure of how much the current strategy is appropriate for achieving the goal. If this affective state is positive, the agent feels confident in the current decision strategy.

Here are two hypotheses on the predicted-utility function whose shape changes with the agent's confidence state. The hypotheses explains how the agent with the confidence-based
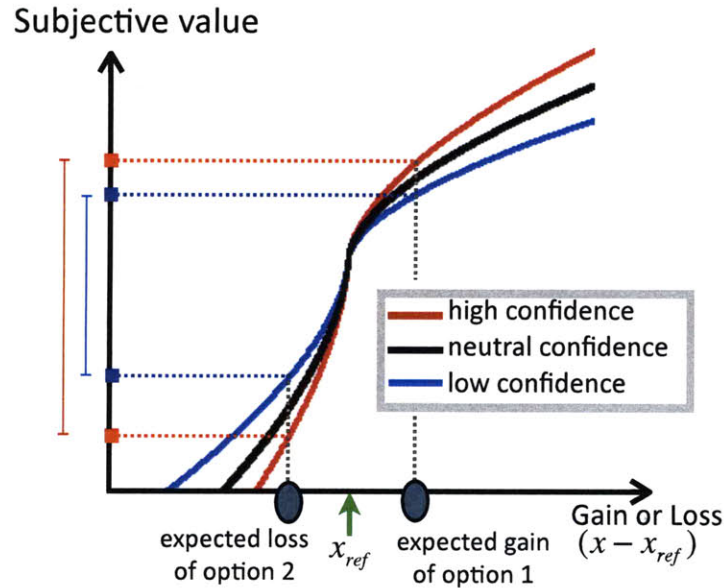
Figure 2-5: The shape of the predicted-utility function changes with the confidence (goal-achieving) state: this change influences discriminability.

affective prediction can be more efficient in controlling the exploitation-exploration balance in learning situations, compared to the agent without the confidence state. Figure 2-5 shows how the confidence-based shape change of the predicted utility function influences discriminability (regarding the separation of expected subjective values of two options).

(1) As the agent becomes more confident in achieving the goal (i.e., a good goal-achieving state), it becomes more sensitive to both likely gains and losses.

(2) As the agent becomes less confident in achieving the goal (i.e., a bad goal-achieving state), it becomes less sensitive to both likely gains and losses.

When the agent feels *high confidence* (good goal-achieving state) from cumulative gains on past choice trials, she has a more sensitive predicted utility function in both domains of gains and losses. Because of this greater sensitivity in high confident state, there are greater separations between the predicted utilities of the currently-estimated best option and the other options. In other words, compared to neutral or low confidence state, the agent in high confidence state feels as if the estimated subjective values of the currently-estimated

best option and the other options were *more* separated. Thus, the agent will have *greater discriminability* and be more likely to select the currently-observed best option (called the *exploitative choice*). On an unpredictable dynamic domain under uncertainty, when the domain is temporarily stationary, the agent will be more likely to make the exploitative choice.

Yet, when the agent feels *low confidence* (bad goal-achieving state) from cumulative losses on past choice trials, she has a less sensitive predicted utility function in both domains of gains and losses. Because of this smaller sensitivity in low confidence state, there are smaller separations between the predicted utilities of the currently-observed best option and the other options. In other words, compared to neutral or high confidence state, the agent in *low confidence* state feels as if the estimated subjective values of the estimated subjective values of the currently-estimated best option and the other options were *less* separated. Thus, the agent will have *smaller discriminability* and be more likely to select currently-observed suboptimal options (called the *exploratory choice*). On an unpredictable dynamic domain under uncertainty, when the domain is temporarily changing, the agent will be more likely to make the exploratory choice.

2.5.7 Incidental emotion state and framing influences experience, prediction and risk attitude

The standard PT subjective value function as shown in Figure is concave ($0 < a < 1$) in the domain of potential gains, and convex ($0 < b < 1$) in the domain of potential losses.

$$v = f_{PT}(x - x_{ref}) = \begin{cases} (x - x_{ref})^a, & x - x_{ref} \geq 0 \\ -\lambda(-(x - x_{ref}))^b, & x - x_{ref} < 0 \end{cases}$$
$$0 < a < 1, \quad 0 < b < 1, \quad \lambda > 1$$

A PT-based parameterized subjective value function can model the risk attitude (risk seeking or risk averse) and the sensitivity either in the domains of potential gains or in the domain of potential losses by changing the values of its parameters.

First, in the domain of potential gains:

- When $a > 1$ (convex): risk-seeking

54

- When $a = 1$ (linear): risk-neutral

- When $0 < a < 1$ (concave): risk-averse

- As a increases, more risk-seeking

- As a decreases, more risk-averse

Second, in the domain of potential losses:

- When $b > 1$ (concave): risk-averse

- When $b = 1$ (linear): risk-neutral

- When $0 < b < 1$ (convex): risk-seeking

- As b increases, more risk-averse

- As b decreases, more risk-seeking

- As λ increases, more sensitive to losses

- As λ decreases, less sensitive to losses

Note that the sensitivities to potential gains and losses, respectively, relate to the asymptotic slopes of the function in the domain of gains and losses, and the uncertainty attitudes in the domain of potential gains and losses relate to the curvatures (convexity and concavity) in the corresponding domain.

In a gambling task involving potential gains, people with a smaller a have more risk-averse attitude, whereas people with a larger a have more risk-seeking attitude. Also, in a gambling task involving potential losses, people with a smaller b (> 0) have more risk-seeking attitude, while people with a larger b have more risk-averse attitude.

The parameter values of the PT-based subjective value function change with the incidental mood state. Based on a variety of psychological experiments on the influences of different mood states on the risk attitude, the hypotheses on the change of parameters under different mood states are as follows.
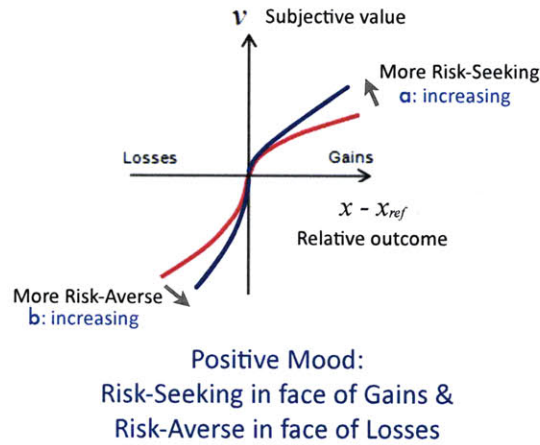
Figure 2-6: Positive mood changes the predicted-utility function

Positive incidental affects

The promoted risk-seeking attitude under positive moods in potential-gain situations [12, 25] could be associated with an increased a (i.e., more risk-seeking and more sensitive to gains). Likewise, the promoted risk-averse attitude under positive moods in potential-loss situations [12, 25] could be associated with an increased b (i.e., more risk-averse and more sensitive to losses). The shape of the subjective value function that Isen and her colleagues [26] experimentally inferred for people in positive moods can be also related to this explanation.

Negative incidental affects

Fear and Anger: The promoted risk-averse attitude in fear [32, 33, 34] could be associated with a decreased a (i.e., more risk-averse and less sensitive to gains) and an increased b (i.e., more risk-averse and more sensitive to losses). Likewise, the promoted risk-seeking attitude in anger [32, 33, 34] could be associated with an increased a (i.e., more risk-seeking and more sensitive to gains) and a decreased b (i.e., more risk-seeking and less sensitive to losses).

(a) Fear (anxiety)   (b) Anger
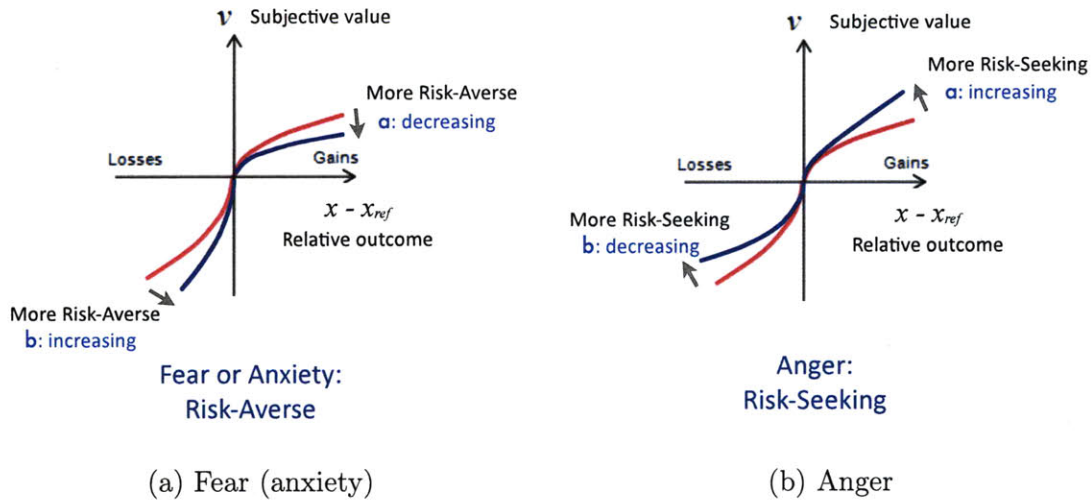
Figure 2-7: Fear (anxiety) and Anger change the predicted-utility function

Anxiety and Sadness: The promoted risk-averse attitude in anxiety [46, 47], which is a fear-like emotion, could be associated with a decreased a (i.e., more risk-averse and less sensitive to gains) and an increased b (i.e., more risk-averse and more sensitive to losses). From the facts that sad people prefer high-risk/high-reward options to low-risk/low-reward options and place a lower value on objects that they have possessed [46, 47], it could be inferred that sad people tend to be less sensitive to potential losses, more sensitive to potential gains and take risks for gains. Also, the reverse endowment effect in sadness [35] could be associated with the less sensitivity to potential losses. This attitude in sad moods could be associated with an increased a (i.e., more risk-seeking and more sensitive to gains) and a decreased λ (i.e., less sensitive to losses).

Disgust: The elimination of endowment effect in disgust [35] could be associated with a decreased λ (i.e., less sensitive to losses).

2.5.8 Comparison with other models

The AC model is distinguished from other cognitive approaches such as Bayesian belief updating [18, 15], model-free reinforcement learning [58], and EWA (experience-weighted attraction) learning [9], in terms of the concept or computation of decision utility. In

(a) Sadness   (b) Disguss
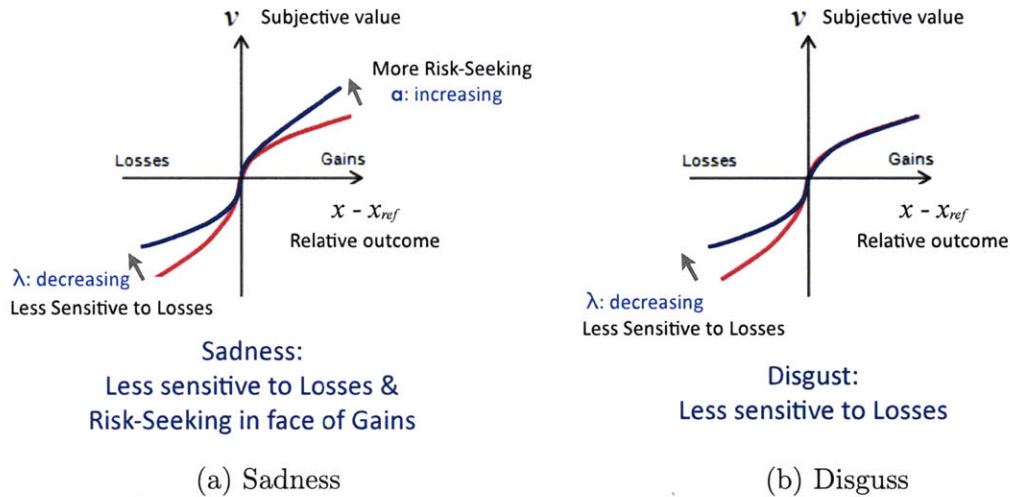
Figure 2-8: Sadness and Disgust change the predicted-utility function

Bayesian belief updating, decision utility for a choice means a belief about future outcomes predicted in Bayesian way. This Bayesian concept of decision utility involves only the cognitive aspect of predicted utility neglecting the affective influence on prediction. In model-free reinforcement learning, decision utility for a choice is a weighted average outcome over past experiences. Thus, decision utility in model-free reinforcement learning represents the cognitive aspect of total-experienced utility neglecting the subjective influence (i.e., reference dependence, loss-aversion and diminishing sensitivity) on outcome evaluations. In EWA learning, decision utility is a combination of the cognitive aspects of predicted utility and total-experienced utility.

In the AC model, affect has two important functional roles. First, affect in the experience phase serves as experience-feedback (reinforcement or punishment) signals. This role of affect is similar to Panksepp's concept of sensory affects [42] or Kahneman's concept of experienced utility [28, 31]. In the AC model, this subjective evaluation signal is modeled by the PT-based parameterized experienced-utility function. Second, affect in the decision-making phase serves as processing modulation signals that influence on both the experience-based mode (total-experienced utility) and the prediction-based mode (predicted utility) as follows:

- In the experience-based mode, a built-in action tendency associated with an affective

state may motivate the decision maker to take an associated action (i.e., an affectively attractive action). The affective state at the time of decision making directly influences the experience-based mode in such a way to automatically reactivate past memories experienced in similar situations to the affective state. In terms of computation modeling, the affective state is used as an index for finding similar situations to be taken into account for the computation of the overall experience value (total-experienced utility). This direct influence is called "affective referral" in the AC model.

- In the prediction-based mode, the affective state influences the computation of the predicted utility through the change of the predicted utility function. Thus, the predicted utility in the AC model is not a pure-cognitive Bayesian belief on future outcomes, but it is affectively modulated belief on the hedonic impacts of future outcomes. This indirect influence is called "affective shaping" in the AC model.

- In the AC model, the decision utility is a linear combination of the total-experienced utility and the predicted utility whose weights can be adjusted by a cognitive-load parameter.

Several important decision-making models in psychology, economics and neuroscience, such as

- Kahneman and Tversky (prospect theory) [30]

- Loewenstein and O'Donoghue (animal spirits: affective and deliberative processes) [37]

- Kahneman (objective happiness: moment-based approach) [28, 31, 29]

- Loewenstein and Lerner (risk-as-feelings, the role of affect in decision making) [38, 36]

- Berridge and Robinson (incentive salience) [7, 6]

- Bechara and Damasio (somatic marker hypothesis) [4, 5]

- Dayan and Daw (Pavlovian and instrumental actions) [16]

- Yechiam and Busemeyer (expectancy-valence model) [63]

- Camerer and Ho (EWA: experience-weighted attraction) [9]

are compared with the AC model in terms of the following criteria:

- What is the main field of the model? (Behavioral economics, Psychology, Neuroscience, Machine learning, etc)

- Is it a model for animal behavior or human decision making? Is the model descriptive or generative or both?

- Does the model provide a computational framework?

- What kinds of situations can the model deal with? (decision making under the full information of risks, or decision making under uncertainty (adaptive learning))

- How does the model evaluate an obtained outcome? What is the concept of experienced utility in the model? Does the model employ a subjective experienced-utility function?

- Does it model the influence of past overall experience on decision making? Does the concept of total-experienced utility involve both cognitive and affective influence?

- Does it model the influence of prediction on decision making? Does the concept of predicted utility involve both cognitive and affective influence? Does the model employ a subjective predicted-utility function?

- Does it model the influence of emotions in decision making?

- Does it model an intertemporal affective dynamics with a model of task-related emotion state?

- How does it explain the decision preference (determining decision behavior)? What kinds of influences contribute to the decision preference?

- What is the rationality criterion of the model?

The following Figures 2-9, 2-10, 2-11 and 2-12 compare the AC model with other models.

	Kahneman & Tversky (Prospect theory (PT))	Kahneman (Objective happiness: Moment-based Approach)	Loewenstein & O'Donoghue (Animal Sprits)	Loewenstein & Lerner (Risk as feelings + Role of affect in DM)	Berridge & Robinson (Incentive salience)	Bechara & Damasio (Somatic marker hypothesis)	Dayan & Daw (Pavlovian and Instrumental actions)	Camerer & Ho (EWA: experience-weighted attraction)	Yechiam & Busemeyer (Expectancy-valence model)	Ahn & Picard (Affective-cognitive (AC) decision model)
Main field	Behavioral Economics, Psychology	Psychology	Behavioral Economics	Psychology	Neuroscience	Psychology, Neuroscience	Neuroscience, Machine learning	Behavioral Economics	Psychology, Neuroscience	Behavioral Economics, Psychology, Neuroscience, Machine learning
Main target	Human Decision Making	Human Decision Making	Human Decision Making	Human Decision Making	Animal Behavior	Human Decision Making	Animal Behavior	Human Decision Making	Human Decision Making	Human Decision Making
Descriptive or Generative	Descriptive	Descriptive	Descriptive	Descriptive	Descriptive	Descriptive	Descriptive, Generative	Descriptive, Generative	Descriptive, Generative	Descriptive, Generative
Math model	Yes	No	Yes	No	No	No	Yes	Yes	Yes	Yes
Decision making under uncertainty	One-shot decision under the full information of underlying risks, Individual DM	.	.	.	.	Multiple-trial adaptive learning situation, Individual DM	Multiple-trial adaptive learning situation, Individual DM	Multiple-trial adaptive learning situation, Multiple players' game	Multiple-trial adaptive learning situation, Individual DM	Multiple-trial adaptive learning situation, Individual DM
How to model the experience of an obtained outcome	Kahneman distinguishes experienced utility from decision utility, but prospect theory itself does not model experienced utility	Experienced utility = Affective hedonic experience (pleasure or displeasure) of the outcome	.	.	Affective or hedonic reactions of the outcome (Liking) Similar to Kahneman's experienced utility and Panksepp's sensory affect	.	The amount of reward (Operationalized payoffs)	The amount of outcome (Operationalized payoffs)	The valence of the payoffs Modeled by a linear combination (loss-aversive) of gains and losses	Experienced utility = affective hedonic experience (pleasure or displeasure) of the outcome Modeled by using a generalized PT subjective value function
How to model the role of past overall experience in decision making	.	Total-experienced utility = Statistical aggregation of past experienced utilities	.	Anticipatory emotion (Integral emotion)	.	Somatic state by primary inducers (innate or learned stimuli that cause pleasurable or aversive states)	Pavlovian conditioning (habitual action) and Instrumental conditioning (goal-directed action)	Choice Reinforcement = Statistical aggregation of past payoffs	Expectancy = Statistical aggregation of past valences	Total-experienced utility (Experience-based mode) = Statistical aggregation of past experienced utilities

61

Figure 2-9: Comparison of different decision-making models in psychology, behavioral economics, neuroscience and machine learning

	Kahneman & Tversky (Prospect theory (PT))	Kahneman (Objective happiness: Moment-based Approach)	Loewenstein & O'Donoghue (Animal Sprits)	Loewenstein & Lerner (Risk as feelings + Role of affect in DM)	Berridge & Robinson (Incentive salience)	Bechara & Damasio (Somatic marker hypothesis)	Dayan & Daw (Pavlovian and Instrumental actions)	Camerer & Ho (EWA: experience-weighted attraction)	Yechiam & Busemeyer (Expectancy-valence model)	Ahn & Picard (Affective-cognitive (AC) decision model)
How to model the role of prediction in decision making	PT subjective value function = Predicted utility	Predicted utility = Cognitive expectation on the future hedonic experience	Deliberative evaluation on the expected value of future outcomes Emotions influence deliberative thinking	Expected (Anticipated) emotion Similar to predicted utility Emotions influence predicted utility	Cognitive wanting = Predicted utility *Note:* Remembered utility (hedonic memory of past experience) is the chief factor determining predicted utility	Reasoning such as a cost-benefit analysis alone is not sufficient for making advantageous decisions	Instrumental conditioning (goal-directed action)	Belief-based expected payoffs	.	Affect-modulated predicted utility (Prediction-based mode) Deliberative estimation + Affective modulation Emotions influence predicted utility Modeled by using a generalized PT subjective value function whose parameters change with emotions
How to model the affective influence in decision making	.	.	Affective evaluation driven by emotions and motivational drives	Anticipatory emotion (genuine visceral reactions and/or emotional feelings that arise from expecting the future consequences of the decision) Incident emotion (emotion that arises from the factors unrelated to decision itself)	Incentive salience 'wanting' (Implicit cue-triggered attraction to salient incentive, depending on mesolimbic dopamine systems) Similar to Panksepp's emotion-action affect (seeking)	Somatic state by secondary inducers (thoughts or memories of primary inducers) Background somatic state: pre-existing somatic state triggered by prior emotional events modulates the sensitization of primary and secondary inductions	Pavlovian value for an action such as approach to cues predicting rewards	.	.	Emotions influence both prediction-based and experience-based modes Anticipatory affect = experience-based mode (the affect index-based retrieval of cached total-experienced utility) Incidental affect and/or incentive salience cues influence the parameters of predicted utility

Figure 2-10: Comparison of different decision-making models in psychology, behavioral economics, neuroscience and machine learning (Continued)

	Kahneman & Tversky (Prospect theory (PT))	Kahneman (Objective happiness: Moment-based Approach)	Loewenstein & O'Donoghue (Animal Sprits)	Loewenstein & Lerner (Risk as feelings + Role of affect in DM)	Berridge & Robinson (Incentive salience)	Bechara & Damasio (Somatic marker hypothesis)	Dayan & Daw (Pavlovian and Instrumental actions)	Camerer & Ho (EWA: experience-weighted attraction)	Yechiam & Busemeyer (Expectancy-valence model)	Ahn & Picard (Affective-cognitive (AC) decision model)
How to model the task-related emotion in learning (inter-temporal affective dynamics)	.	.	.	.	.	.	.	.	.	Confidence state (Goal-achieving state) influences prediction and discriminability
How to model the decision preference (determining decision behavior)	PT subjective value function (= Predicted utility)	Decision utility (≠ Experienced utility) (≠ Predicted utility)	Objective = Deliberative utility + Affective influence	Decision = Expected affect + Immediate (anticipatory or incidental) affect	Decision utility = Cognitive wanting (predicted utility) + Incentive salience 'wanting'	Overall somatic state	Action value = Instrumental (goal-directed) value + Pavlovian (habitual) value	Attraction = Choice Reinforcement + Belief-based expected payoffs	Expectancy	Decision utility = Prediction-based mode (Affect-modulated Predicted utility) + Experience-based mode (Total-experienced utility)
Rationality Criterion	.	To maximize the future experienced utility People are bounded rational	Welfare measures discussed Candidate welfare measures: the deliberative system's utility function or the objective function	.	To make decisions according to cognitive wanting (i.e., rational when decision utility is the same as predicted utility)	.	To make decisions according to instrumental (goal-directed) value	.	.	To maximize the future experienced utility People are bounded rational Both modes (experience-based, prediction-based) may help or harm optimal decisions under uncertainty, frames and emotions

Figure 2-11: Comparison of different decision-making models in psychology, behavioral economics, neuroscience and machine learning (Continued)

	Kahneman & Tversky (Prospect theory (PT))	Kahneman (Objective happiness: Moment-based Approach)	Loewenstein & O'Donoghue (Animal Sprits)	Loewenstein & Lerner (Risk as feelings + Role of affect in DM)	Berridge & Robinson (Incentive salience)	Bechara & Damasio (Somatic marker hypothesis)	Dayan & Daw (Pavlovian and Instrumental actions)	Camerer & Ho (EWA: experience-weighted attraction)	Yechiam & Busemeyer (Expectancy-valence model)	Ahn & Picard (Affective-cognitive (AC) decision model)
Notes	The probability weighting function was modeled as inverse-S shape low probability (near zero): overweighted high probability (near one): underweighted	Accurate prediction of future tastes (accurate predicted utility) and accurate evaluation of past experiences (accurate remembered utility) emerge as critical elements of an individual's ability to maximize the experienced quality of his outcomes	A person's behavior is the outcome of an interaction between two systems: a *deliberative* system and an *affective system* This two-system view explains many departures from full rationality in human behaviors, and captures the familiar feeling of being "of two minds."	Affect typically plays an informational role in decision making. However, "risk-as-feelings" hypothesis posits that, in addition, anticipatory emotional reactions sometimes diverge from cognitive evaluations	A choice is rational so long as it maximizes both decision utility and predicted utility Decision rationality cannot be held responsible for the eventual unhappy experienced utility, because rationality in this sense cannot be held accountable for the accuracy of predictions	Knowledge and reasoning alone are usually not sufficient for making advantageous decisions, and the role of emotion in decision-making has been underestimated Emotion is beneficial to decision-making when it is integral to the task, but can be disruptive when it is unrelated to the task	Competition between goal-directed and habitual controls rests on inferences about their relative accuracies A great deal will turn on the tradeoff between the computational simplicity of Pavlovian control and the danger of suboptimal misbehavior	.	.	True rationality increases with more time and resources, more considered inputs, more accurately evaluated experience, and acknowledge of the current emotion state and its influences

Figure 2-12: Comparison of different decision-making models in psychology, behavioral economics, neuroscience and machine learning (Continued)

Chapter 3

Modeling Subjective Experience and Affective Prediction

3.1 Utilities and subjective experience

Kahneman [28, 31] suggests that there are distinct concepts of utility: the pleasure or displeasure of subjective experience of a choice outcome is defined as "experienced utility", that is, the affective and hedonic concept of utility, whereas the probability of selecting a choice draws on "decision utility", that is, the behavioral and motivational concept of utility. Thus, decision utility is the concept used for describing actual (observed) choice behaviors and this concept is similar to that of option value in the reinforcement-learning literature. Also, "predicted utility" is the decision-maker's belief about the future experienced utility of a choice outcome.

The role of subjective prediction in decisions under risk – when outcome probabilities of each choice are explicitly described in the problem and fully known to the decision maker – has been extensively examined in traditional decision theories such as prospect theory. For decisions under risk, the main determinant of decisions is the decision maker's predicted utility (i.e., decision utility = predicted utility). Prospect theory employs a subjective utility function called the "predicted-utility function" by which the decision maker's risk attitudes and framing (dependence on the reference point) can be described for decisions under risk.

Yet, the role of subjective experience in decisions under uncertainty – when outcome

probabilities of each choice are not explicitly described and should be learned from past experiences by the decision maker – has less been investigated, compared to the role of subjective prediction in decisions under risk. For decisions under uncertainty, the overall experience of the decision maker on previous trials in the same situation has a critical impact on future decisions, the experience-based mode (total-experienced utility) critically influences decision utility.

In this section I focus on two-armed bandit problems with stationary distributions of stochastic outcomes to understand the impacts of past overall subjective experience on the current decision through both computational approaches. To model subjective experience, I proposed and tested another kind of a PT-based parameterized subjective value function called the "experienced-utility function", which is independent of and separate from the predicted-utility function. I investigate how framing (reference point selection) influences subjective experiences and in turn, the "(subjective) discriminability" of choices. The concept of discriminability characterizes the level of easiness in figuring out which choice has a greater average utility than other ones with fewer trials; thus, the discriminability is a key factor in regulating the trade-offs between exploration and exploitation and quickly detecting the optimal decision in learning. Computational simulation results will show that discriminability can be increased by the use of the experienced-utility function in some domains with an appropriate reference point.

The nonlinear shape of the experienced-utility function reflects the decision maker's risk-averse experience when it comes to their gains, and risk-seeking experience when it comes to their losses. This implies that a distinct shape of the experienced-utility function may lead to a different "subjective discriminability" for the same underlying outcome distribution. An increased subjective discriminability enables the decision maker to need fewer exploratory choice trials to achieve the best overall outcome. That is, the decision maker when employing a nonlinear experienced-utility function with a good decision frame (or a good reference point) may achieve greater overall outcome than when employing a linear experienced-utility function. Note that the decision maker with a linear experienced-utility function has the same discriminability (or "objective discriminability") regardless of the reference point selection.

Although reinforcement-learning (RL) algorithms focus on decisions under uncertainty, most have not focused on how the shape of the function modeling subjective experiences influences the decision maker's learning and discriminability; in most applications, the experienced-utility function is simply assumed to be linear.

While PT theory has been developed in economics for domains with known outcome distributions, this new model enables PT theory to be used for unknown and changing stochastic outcome distributions.

3.2   Decisions under uncertainty

In the decision-making literatures, decision situations under "risk" (stochastic outcomes but known outcome distributions of alternatives) are distinguished from decision situations under "uncertainty" (stochastic outcomes and unknown outcome distributions of alternatives). Thus, in situations under uncertainty, decision makers should learn the values of each option from their overall previous experience.

In this section, I focus on the experience-based mode in the AC model and investigate the benefits and pitfalls that it can bring in different decision-making domains under uncertainty.

The experience-based mode is similar to the model-free RL, but it employs an "experienced-utility function" to evaluate the subjective value (called "experienced utility" or "moment utility") of the obtained outcome and update the cached overall subjective value (called "total-experienced utility") of the choice by current experienced utility. In this way, the experience-based mode can model the decision maker's different risk attitude (risk-seeking or risk-averse with experienced gains and losses) and sensitivity to experienced losses. Note that "predicted-utility function" models the decision maker's risk attitude in face of potential gains and losses and sensitivity to potential losses.

Here the experience-based mode of the AC agent employing the PT-based experienced-utility function (i.e., experienced utility = subjective value (pleasure or displeasure)) will be compared with the model-free RL agent employing the linear utility function (i.e., experienced utility = objective outcome (gain or loss)).

The total-experienced utility (or weighted average over past sampled experienced utilities) for an option represents the agents overall experienced feeling (i.e., positive or negative valence) about that option. Assume that the AC agent only depends on the experience-based mode (i.e., decision utility = total-experienced utility). Under this assumption, if option 1 has a greater total-experienced utility than option 2, the difference in the utilities for two options shows the relative preference in choosing option 1 over option 2. In this case, option 1 is called "positively valenced" and option 2 is called "negatively valenced", where these labels reflect relative preference.

Two distinct kinds of stationary two-armed bandit domains are considered to explain how subjective valuation in the experience phase influence the earlier detection of the optimal option in some domain.

In two-armed bandit problems, there are two options (i.e., option 1 and option 2). Thus, when option k (=1,2) is selected and then outcome x is obtained, the total-experienced utility $V(k)$ is updated as follows:

(1) Reference point x_{ref} is modeled by the average of the estimated expected outcomes of option 1 and option 2: $x_{ref} = (\hat{\mu}(1) + \hat{\mu}(2))/2$

Note that there could be other ways of reference point selection. This average-based reference is one particular setup for simulations here.

(2) Experienced utility v for outcome x and reference point x_{ref} is computed through the experience-utility function f_{EU} in the current state: $v = f_{EU}(x - x_{ref})$.

(3) Total-experienced utility for option k is updated: $V(k) \leftarrow V(k) + \alpha \ (v - V(k))$ where α is the learning rate parameter.

(4) The estimated expected outcome for option k is updated as follows (if it is not separately modeled in the prediction-based mode): $\hat{\mu}(k) \leftarrow \hat{\mu}(k) + \alpha \ (x - \hat{\mu}(k))$ where α is the learning rate parameter.

Note that when $\alpha = count(k)$ where $count(k)$ is the number of trials of option k until now, $\hat{\mu}(k)$ and $V(k)$, respectively, compute the exact averages of sampled outcomes and sampled subjective values of option k. In the simulations on the stationary domains below, $\alpha = count(k)$ is assumed.
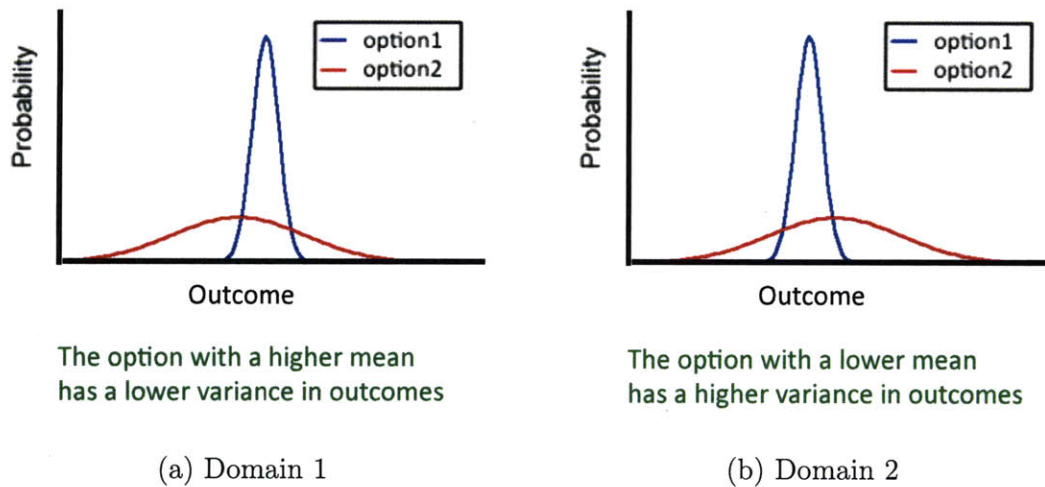
(a) Domain 1   (b) Domain 2

Figure 3-1: (Example) domains under uncertainty: Domain 1 vs. Domain 2

3.2.1 Objective discriminability vs. Subjective discriminability

Task Domains: The following two-armed bandit tasks (See Figure 3-1) used in simulations are very useful in examining how framing influences subjective experience and optimal decisions.

1. There are two options: one "risky" option with high variance in outcomes and the other "safe" option with low variance in outcomes. Both options have Gaussian outcome distributions.

2. There were two domains (Domain 1 and Domain 2) used in simulations: For Domain 1, the safe option is the optimal option (i.e.,the option with a greater outcome on average) and the risky option is the suboptimal option (or with a smaller outcome on average). For Domain 2, the risky option is the optimal option and the safe option is sub-optimal.

Discriminability: The concept of discriminability has been largely investigated under different names in a variety of areas such as psychophysical judgment and decision theory [60, 8], pattern classification [18], signal detection theory (called the "sensitivity index" or d') [62] and statistical power analysis (called the "effect size") [11]. Discriminability can be used for characterizing the level of easiness in discriminating which choice is optimal with fewer trials. Figure 3-2 shows how the "true" objective and subjective discriminabilities are

defined. The subjective value distributions are the transformation of the objective outcome distributions through a PT-based subjective value function (experienced-utility function). Note that the transformed subjective value distributions (and the subjective discriminability) depend on the reference point location as well as the shape of the experienced-utility function. The subjective discriminability depends on the decision maker's subjective-value distributions, whereas the objective discriminability depends on the original outcome distributions.

A larger subjective discriminability makes the decision maker more likely to choose the optimal option on the next trial, whereas a smaller subjective discriminability makes the decision maker more likely to choose the suboptimal option. Also, a positive value of subjective discriminability indicates that experiences contribute to selecting the optimal option more, whereas a negative value means the opposite tendency.

If the subjective discriminability is greater than the objective discriminability, the agent with a subjective value function can tell which option is optimal with fewer exploratory trials, compared to the agent with a linear utility function.

Below are the definitions of the "estimated" objective discriminability and subjective discriminability in a two-armed bandit task.

Defining objective discriminability: d'_{obj} (objective discriminability) is defined as

$$d'_{obj} = \frac{\hat{\mu}_1 - \hat{\mu}_2}{\sqrt{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}}$$

where
$\hat{\mu}_1 = $ the average of obtained outcomes from the optimal option (i.e., the safe option on Domain 1 or the risky option on Domain 2) over trials,
$\hat{\mu}_2 = $ the average of obtained outcomes from suboptimal option (i.e., the risky option on Domain 1 or the safe option on Domain 2) over trials,
$\hat{\sigma}_1 = $ the standard deviation of obtained outcomes from the optimal option over trials,
$\hat{\sigma}_2 = $ the standard deviation of obtained outcomes from the suboptimal option over trials.

The definition of objective discriminability depends on a decision maker's obtained out-
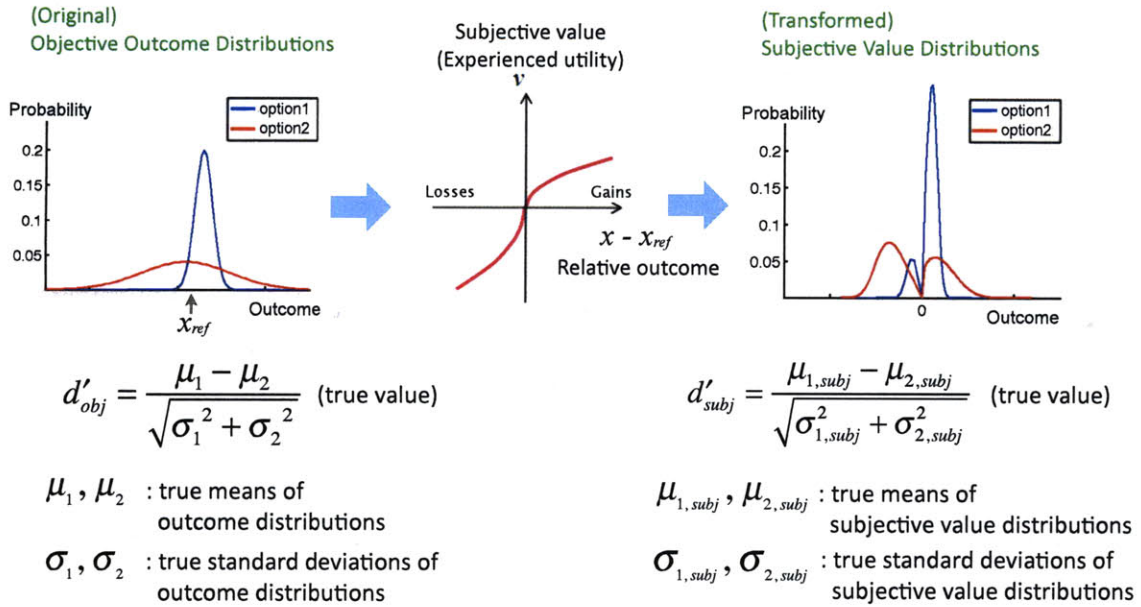
Figure 3-2: Objective discriminability d'_{obj} vs. Subjective discriminability d'_{subj} (true values)

comes over trials. That is, the measure of objective discriminability relies only on the underlying outcome distributions and samplings from them, but should not be affected by the gain or loss frame used to evaluate experienced utilities.

Defining subjective discriminability: d'_{subj} (subjective discriminability) is defined as

$$d'_{subj} = \frac{\hat{\mu}_{1,subj} - \hat{\mu}_{2,subj}}{\sqrt{\hat{\sigma}^2_{1,subj} + \hat{\sigma}^2_{2,subj}}}$$

where

$\hat{\mu}_{1,subj}$ = the average of experienced-utilities from the optimal option over trials,

$\hat{\mu}_{2,subj}$ = the average of experienced-utilities from suboptimal option over trials,

$\hat{\sigma}_{1,subj}$ = the standard deviation of experienced-utilities from the optimal option over trials,

$\hat{\sigma}_{2,subj}$ = the standard deviation of experienced-utilities from the suboptimal option over trials.

The following sections will show why the above definitions of discriminabilities are useful

71

in the analysis of optimal decision making under uncertainty: the concept of discriminability is associated with the probability of selecting the optimal option after some exploratory trials (assuming the greedy selection strategy). That is, a higher discriminability leads to a greater probability of selecting the optimal option.

3.2.2 Objective discriminability

Consider a two-armed bandit task in which each option k (=1, 2) is associated with a normal (Gaussian) outcome distribution $r \sim \mathrm{N}(\mu_k, \sigma_k^2)$ (assuming $\mu_1 > \mu_2$). Note that these underlying outcome distributions are initially unknown to the decision maker. The goal of the decision maker is to maximize the total outcome during N trials. It is well known that decision problems under uncertainty require the dilemma between exploration and exploitation [59, 39]. Most strategies to solve this dilemma have proposed some initial exploratory trials or the exploration bonus. To easily define the relevant concepts of discriminability associated with exploration and exploitation, here I focus on the greedy selection strategy with initial $2n_B$ exploratory trials (assuming $2n_B << N$). In this strategy, the decision maker clearly distinguishes initial $2n_B$ exploratory trials from later $N - 2n_B$ exploitative trials. During the exploratory trials, the decision maker alternatively selects one of the options; thus, after these trials, random outcomes of n_B trials for each option will be obtained. During the exploitative trials, the decision maker selects the option with a current higher average outcome (or sample mean of obtained outcomes during past trials from the option): denoting the average outcome of each option as $\hat{\mu}_1^t$ and $\hat{\mu}_2^t$ on trial t ($\geq t_B \triangleq 2n_B + 1$), the decision maker selects option 1 when $\hat{\mu}_1^t > \hat{\mu}_2^t$ and option 2 when $\hat{\mu}_1^t < \hat{\mu}_2^t$.

Here, to define a concept of discriminability associated with the initial $2n_B$-trial exploration, I focus on the trial $t_B (= 2n_B + 1)$ immediately after $2n_B$ exploratory trials. On this trial the average outcome or the sample mean of n_B observed outcomes after n_B exploratory trials of each option $k(= 1, 2)$ is computed as $\hat{\mu}_k^{t_B} \triangleq (1/n_B) \sum_{i=1}^{n_B} r_k^{(i)}$ where $r_k^{(i)}$ is the ith sampled outcome of option k. Also, sample means $\hat{\mu}_k^{t_B}$ follow normal distributions: $\hat{\mu}_k^{t_B} \sim \mathrm{N}(\mu_k, (\sigma_k/\sqrt{n_B})^2)$ for each k. Now the probability of choosing option 1 over option 2 on trial t_B ($= 2n_B + 1$) according to the average objective outcomes is $\mathrm{Pr}_{obj}(\text{option} = 1) = \mathrm{Pr}(\hat{\mu}_1^{t_B} > \hat{\mu}_2^{t_B}) = \mathrm{Pr}(\hat{\mu}_1^{t_B} - \hat{\mu}_2^{t_B} > 0) = \mathrm{Pr}(y > 0)$ where

$y \triangleq \hat{\mu}_1^{t_B} - \hat{\mu}_2^{t_B}$. Since $\hat{\mu}_1^{t_B}$ and $\hat{\mu}_2^{t_B}$ are normal variables, y is also a normal variable following $y \sim N(\mu_1 - \mu_2, (\sigma_1^2 + \sigma_2^2)/n_B)$. Now the standard normal variable $z = \frac{y-(\mu_1-\mu_2)}{\sqrt{(\sigma_1^2+\sigma_2^2)/n_B}} \sim N(0,1)$ whose cumulative distribution function (cdf) is $\Phi(x) = \frac{1}{2}\left(1 + erf(\frac{x}{\sqrt{2}})\right)$ leads to $\Pr(y > 0) = \Pr(z > -d_B) = 1 - \Phi(-d_B) = \Phi(d_B)$ where $d_B = \frac{\mu_1-\mu_2}{\sqrt{(\sigma_1^2+\sigma_2^2)/n_B}}$.

Defining the *objective* discriminability (called *objective d-prime*) $d'_{obj} \triangleq \frac{\mu_1-\mu_2}{\sqrt{(\sigma_1^2+\sigma_2^2)}}$, $d_B = \sqrt{n_B}\, d'_{obj}$ and thus, $\Pr_{obj}(\text{option} = 1) = \Phi(\sqrt{n_B}\, d'_{obj})$. Note that d'_{obj} depends only on the statistics of objective outcome distributions given in the problem and that as d'_{obj} of the underlying domain increases, the *objective* decision maker's probability of choosing option 1 over option 2 after $2n_B$ exploratory trials becomes close to 1 with the objective-outcome based greedy selection strategy.

3.2.3 Subjective discriminability

Now consider what happens to the discriminability when the decision maker employs the subjective value (or experienced-utility) function with a pre-fixed reference point.

Monte Carlo simulations can be used to estimate the statistics of the *subjective value* distributions ($v_k = f(r_k)$ for $k = 1, 2$) obtained by shaping the original *objective outcome* distributions ($r_k \sim N(\mu_k, \sigma_k^2)$) through the subjective value function $f(\cdot)$. The true means ($\mu_{subj,k}$) and standard deviations ($\sigma_{subj,k}$) of the *subjective value* distributions can be estimated as follows: for a large positive integer M and the mth sampled outcome $r_k^{(m)} \sim N(\mu_k, \sigma_k^2)$ ($m = 1, \cdots, M$),

$$\mu_{subj,k} = E[v_k] = E[f(r_k)] = (1/M)\sum_{m=1}^{M} f(r_k^{(m)})$$

$$\sigma_{subj,k} = \sqrt{Var[v_k]} = \sqrt{Var[f(r_k)]} = \sqrt{1/(M-1)\sum_{m=1}^{M}(f(r_k^{(m)}) - \mu_{subj,k})^2} \quad .$$

The average subjective value of option k after n_B exploratory trials is the sample mean of n_B subjective values, $\hat{\mu}_{subj,k}^{t_B} \triangleq (1/n_B)\sum_{i=1}^{n_B} v_k^{(i)}$. Assuming the number of exploratory trials (n_B) is sufficiently large, according to the central limit theorem, I approximate the distributions of the subjective-value sample means $\hat{\mu}_{subj,k}^{t_B}$ by normal distributions: $\hat{\mu}_{subj,k}^{t_B} \sim N(\mu_{subj,k}, (\sigma_{subj,k}/\sqrt{n_B})^2)$ for option $k(= 1, 2)$. Now the probability of choosing option 1

over option 2 on trial $t_B(= 2n_B + 1)$ according to the average subjective values $\hat{\mu}^{t_B}_{subj,1}$ and $\hat{\mu}^{t_B}_{subj,2}$ can be obtained in a manner similar to computing the probability according to the average outcomes. $\Pr_{subj}(\text{option} = 1) = \Pr(\hat{\mu}^{t_B}_{subj,1} > \hat{\mu}^{t_B}_{subj,2}) = \Pr(\hat{\mu}^{t_B}_{subj,1} - \hat{\mu}^{t_B}_{subj,2} > 0) = \Phi(d_{subj,B})$ where $d_{subj,B} = \frac{\mu_{subj,1} - \mu_{subj,2}}{\sqrt{(\sigma^2_{subj,1} + \sigma^2_{subj,2})/n_B}}$.

Defining the *subjective* discriminability (called *subjective d-prime*) $d'_{subj} \triangleq \frac{\mu_{subj,1} - \mu_{subj,2}}{\sqrt{(\sigma^2_{subj,1} + \sigma^2_{subj,2})}}$, the following relationships are obtained: $d_{subj,B} = \sqrt{n_B}\, d'_{subj}$ and $\Pr_{subj}(\text{option} = 1) = \Phi(\sqrt{n_B}\, d'_{subj})$. Note that d'_{subj} depends not only on the underlying outcome distributions, but also on the decision maker's subjective value function whose shape and reference point are described by the parameters. As d'_{subj} increases, the subjective decision maker's probability of choosing option 1 over option 2 after $2n_B$ exploratory trials becomes close to 1 with the subjective-value based greedy selection strategy.

3.2.4 Comparison between objective and subjective discriminabilities

The decision maker's probability of choosing option 1 over option 2 after n_B trials of each option depends on their discriminability (d'_{obj} or d'_{subj}): $\Pr_{obj}(\text{option} = 1) = \Phi(\sqrt{n_B}\, d'_{obj})$ and $\Pr_{subj}(\text{option} = 1) = \Phi(\sqrt{n_B}\, d'_{subj})$. Therefore, if subjective discriminability d'_{subj} is greater than objective discriminability d'_{obj} for a decision maker with an appropriate shape and reference point of the subjective value function in some domains, subjective decision making may provide a better overall performance (or the rate of choosing the optimal option during all trials) due to a higher probability of choosing option 1 over option 2 during the remaining exploitative trials. In other words, to get the same level of confidence of which option is optimal (which is proportional to $\Pr_{obj}(\text{option} = 1)$ or $\Pr_{subj}(\text{option} = 1)$), subjective decision making with a larger d'_{subj} should usually require fewer exploratory trials than objective decision making with a smaller d'_{obj} .

3.2.5 Interaction effects of domain and framing on subjective discriminability

Figure 3-3 graphically represents how the domain type (Domain 1 and Domain 2) and the framing (gain frame or loss frame) influence subjective discriminability. Here the gain frame or loss frame, respectively, is defined as the frame viewing sampled outcomes from

(a) Domain 1, Loss frame

(b) Domain 1, Gain frame

(c) Domain 2, Loss frame

(d) Domain 2, Gain frame
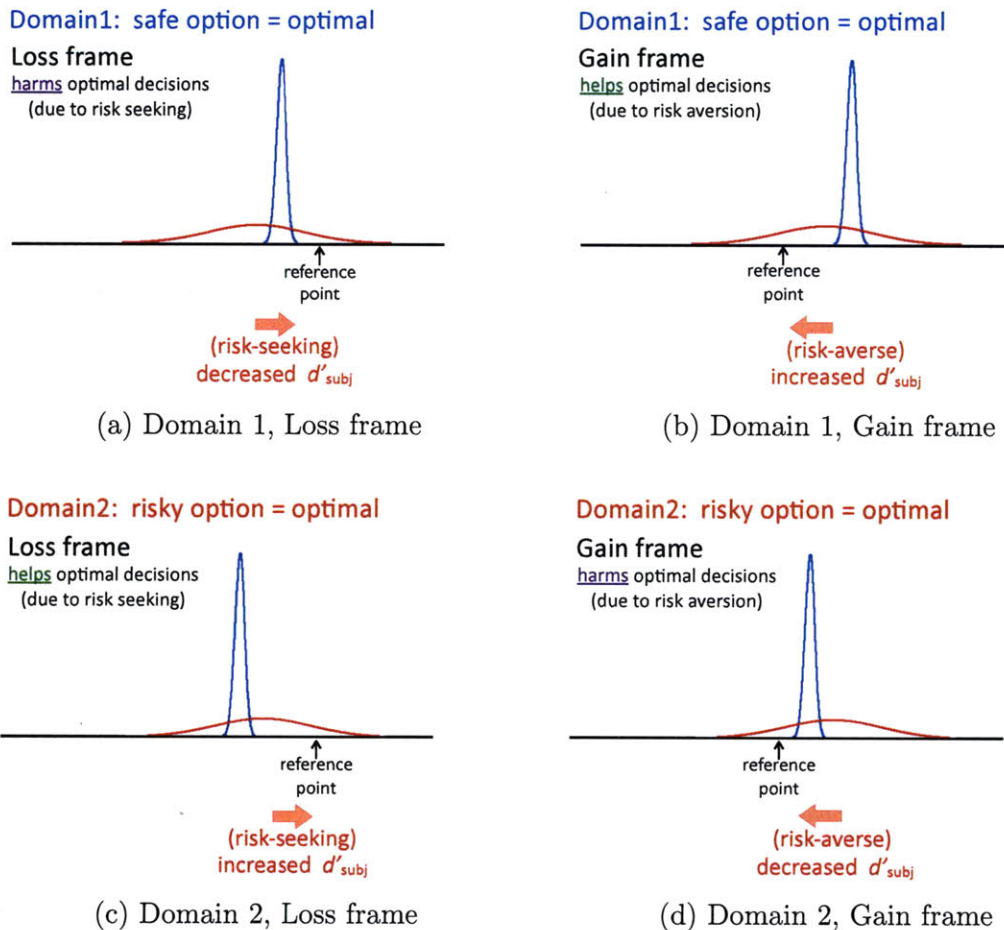
Figure 3-3: The interaction of domain and framing influences subjective discriminability and optimal decisions

the *safe* option as gains or losses. That is, the reference point location determines whether the framing is the gain or loss frame. By the definition of subjective discriminability, the increased d'_{subj} leads to more selections of the optimal option (the safe option on Domain 1, the risky option on Domain 2).

- On Domain 1 (where optimal option = safe option) (Figure 3-3 (a) and (b)), the gain frame increases d'_{subj} due to the risk-averse attitude (motivating to select the safe option), whereas the loss frame decreases d'_{subj} due to the risk-seeking attitude (motivating to select the risky option). Yet, the framing does not influence d'_{obj}.

- On Domain 1, for the decision maker with the PT-based subjective value function, the

75

gain frame helps optimal decision making but the loss frame harms optimal decision making.

- On Domain 2 (where optimal option = risky option) (Figure 3-3 (c) and (d)), the gain frame decreases d'_{subj} due to the risk-averse attitude (motivating to select the safe option), whereas the loss frame increases d'_{subj} due to the risk-seeking attitude (motivating to select the risky option). Yet, the framing does not influence d'_{obj}.

- On Domain 2, for the decision maker with the PT-based subjective value function, the loss frame helps optimal decision making but the gain frame harms optimal decision making.

3.2.6 The influence of the reference point selection (framing) on the subjective discriminability

Figures 3-4, 3-5 and 3-6 show the simulation results on how the reference point selection (framing) influences subjective discriminability on different domains (Domain 1, Domain 2, and a domain where two options have equal variance in outcomes, respectively) for the decision maker employing a subjective value function (experienced-utility (EU)) function with shape parameters $a = 0.8, b = 0.5, \lambda = 2.5$. In all simulations, the estimated discriminabilities d'_{subj} and d'_{obj} were computed based on 10000 randomly sampled outcomes from distributions of each option. Note that d'_{subj} significantly changes as the reference point location changes, whereas d'_{obj} (estimate) is almost the same as the true objective discriminability (d'_{obj} (true)). This is because objective discriminability does not depend on the reference point selection.

On Domain 1 (Figure 3-4), option 1 (average outcome $\mu_1 = 5$ and standard deviation σ_1 = 5) is safe (lower variance) and optimal (higher average), and option 2 (average outcome $\mu_1 = -5$ and standard deviation $\sigma_1 = 10$) is risky (higher variance) and suboptimal (lower average). In this domain, the gain frame (e.g., $x_{ref} = -5$) and the neutral frame, whose reference point is near the mean of the average outcomes of two options, e.g., $x_{ref} = 0$, lead to an increased subjective discriminability ($d'_{subj} > d'_{obj}$), whereas the loss frame (e.g., $x_{ref} \geq 5$) leads to a decreased subjective discriminability ($d'_{subj} < d'_{obj}$).
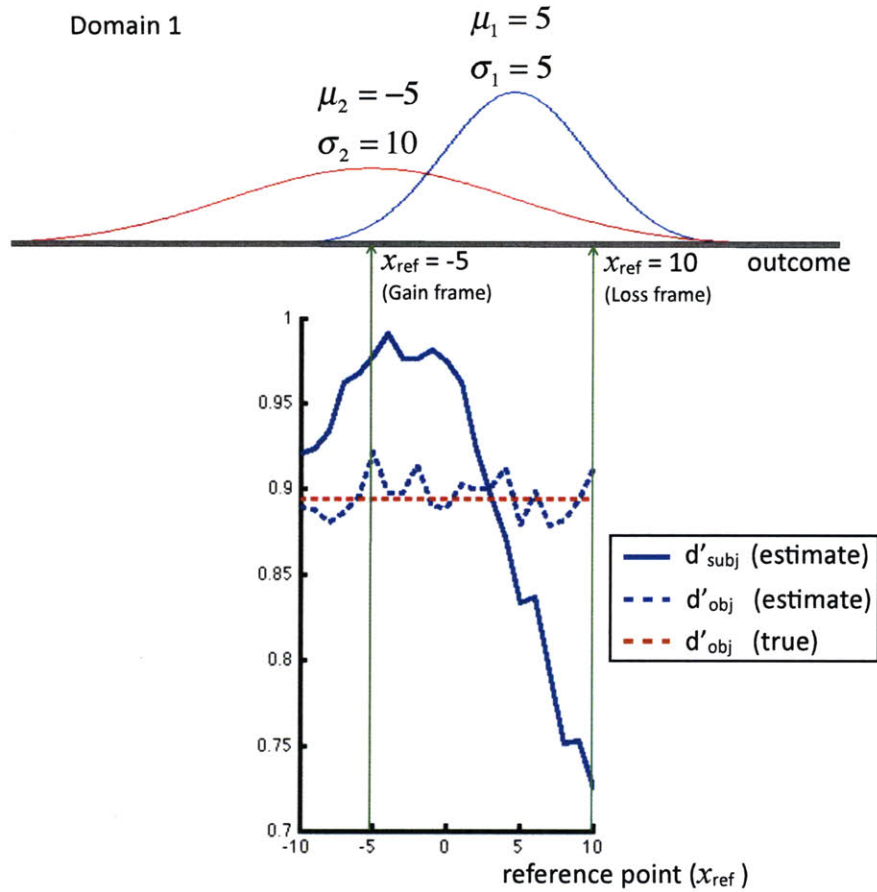
Figure 3-4: Discriminabilities vs. reference point (Domain 1): showing how the reference point selection influences d'_{subj}, employing the EU function $a = 0.8, b = 0.5, \lambda = 2.5$. Green lines indicate two example reference-point selections to show the framing effect.

On Domain 2 (Figure 3-5), option 1 ($\mu_1 = 5$, $\sigma_1 = 10$) is risky (higher variance) and optimal (higher average), and option 2 ($\mu_1 = -5$, $\sigma_1 = 5$) is safe (lower variance) and suboptimal (lower average). In this domain, the loss frame (e.g., $x_{ref} = 5$) and the neutral frame, whose reference point in near the mean of the average outcomes of two options, e.g., $x_{ref} = 0$, lead to an increased subjective discriminability ($d'_{subj} > d'_{obj}$), whereas the gain frame (e.g., $x_{ref} \leq -5$) leads to a decreased subjective discriminability ($d'_{subj} < d'_{obj}$).

On the domain with equal variance (Figure 3-6), option 1 (average outcome $\mu_1 = 5$ and standard deviation $\sigma_1 = 5$) is optimal (higher average), and option 2 (average outcome $\mu_1 = -5$ and standard deviation $\sigma_1 = 5$) is suboptimal (lower average). In this domain,
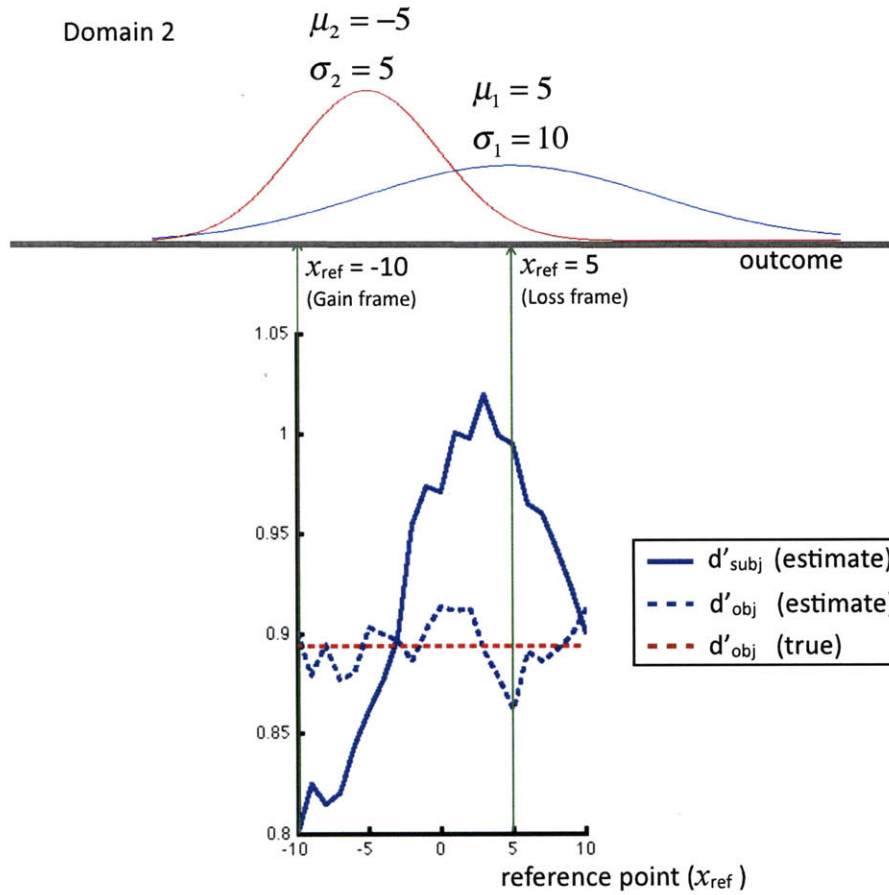
Figure 3-5: Discriminabilities vs. reference point (Domain 2): showing how the reference point selection influences d'_{subj}, employing the EU function $a = 0.8, b = 0.5, \lambda = 2.5$. Green lines indicate two example reference-point selections to show the framing effect.

the neutral frame, whose reference point is near the mean of the average outcomes of two options, e.g., $x_{ref} = 0$, leads to an increased subjective discriminability $(d'_{subj} > d'_{obj})$, whereas gain and loss frames (e.g., $x_{ref} = -10$ (gain frame), $x_{ref} = 10$ (loss frame)) lead to a decreased subjective discriminability $(d'_{subj} < d'_{obj})$.

From these simulations (Figures 3-4, 3-5 and 3-6) on how the reference point selection influences the subjective discriminability on different domains, it is important to note that the neutral frame (whose reference point is near the mean of the average outcomes of two options) led to an increased subjective discriminability independently of underlying domains (Domain 1, Domain 2, Domain with equal variance). In other words, the neutral frame
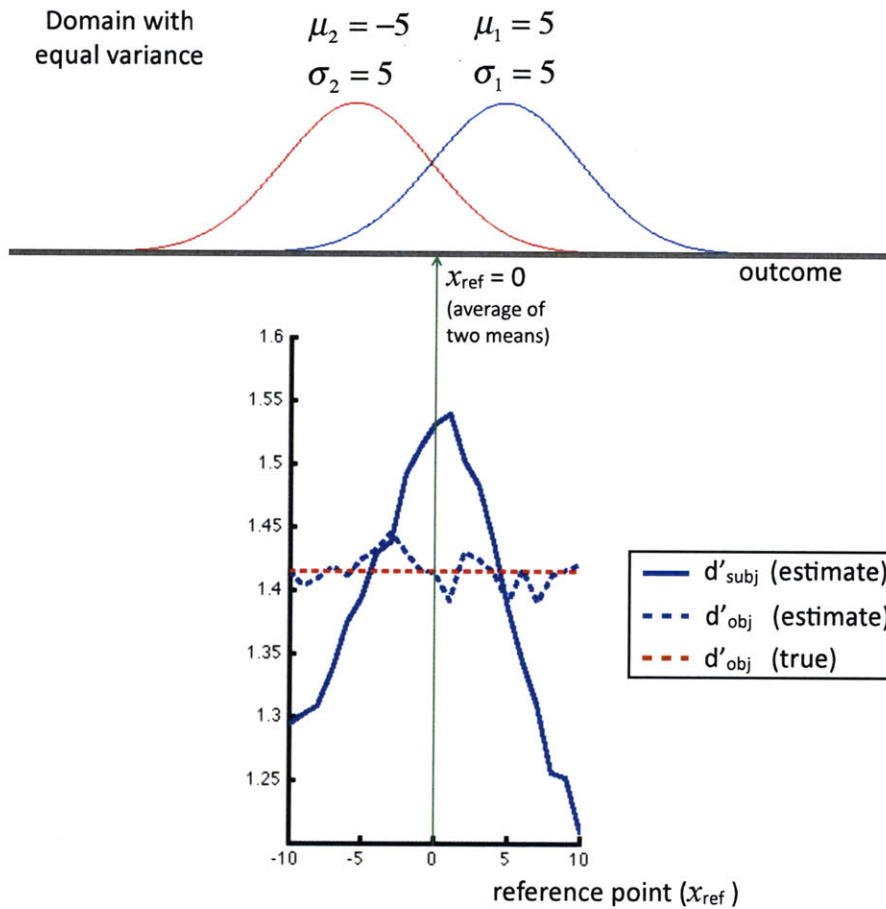
Figure 3-6: Discriminabilities vs. reference point (on a domain where two options have equal variance in outcomes), employing the EU function $a = 0.8, b = 0.5, \lambda = 2.5$. Green lines indicate two example reference-point selections to show the framing effect.

helped make more optimal decisions in learning independently the underlying domains. Thus, it may be beneficial for the decision maker to employ a reference point near the mean of the average outcomes of two options (i.e., the currently-estimated best and second-best options) leading to more optimal decisions in learning under uncertainty.

Similarly, when the reference point is near the mean outcome of the risky option (option 2 on Domain 1, option 1 on Domain 2), the decision maker obtained an increased subjective discriminability. In other words, the gain frame on Domain 1 (e.g., $x_{ref} = -5$) and the loss frame on Domain 2 (e.g., $x_{ref} = 5$) led to an increased subjective discriminability, and those reference point selections corresponded to the mean outcome of the risky option on

each domain. Thus, it may be also beneficial for the decision maker to employ a reference point near the mean of the risky option.

3.2.7 The influence of outcome variances on discriminabilities

Figure 3-7 shows how the outcome variances of two options (σ_1 and σ_2) influence discriminabilities when the decision maker employs different subjective value functions.

First, subplots (a) and (c) in Figure 3-7 show simulation results on Domain 1 where $\mu_1 - \mu_2 = 10$ (fixed), σ_1 is varying from 1 to 5, and $\sigma_2 = 2\sigma_1$. Note that subplots (a) and (c) employed different sets of subjective value function parameters. Here the parameter set employed for subplot (c) ($a = 0.5, b = 0.4, \lambda = 2.5, x_{ref} = 1$) led to greater subjective discriminability over varying σ_1, compared to that employed for subplot (a) ($a = 0.8, b = 0.5, \lambda = 2.5, x_{ref} = 0$).

Second, subplots (b) and (d) in Figure 3-7 show simulation results on Domain 2 where $\mu_1 - \mu_2 = 10$ (fixed), σ_1 is varying from 2 to 10, and $\sigma_2 = 0.5\sigma_1$. Note that subplots (b) and (d) employed different sets of subjective value function parameters. Here the parameter set employed for subplot (d) ($a = 0.5, b = 0.4, \lambda = 2.5, x_{ref} = 1$) led to greater subjective discriminability over varying σ_1, compared to that employed for subplot (b) ($a = 0.8, b = 0.5, \lambda = 2.5, x_{ref} = 0$).

On both domains the subjective discriminability is reliably greater than the objective discriminability when the levels of outcome variances of each option are not very large. Also, the subjective value function parameters (the shape and the reference point) influence the subjective discriminability. Thus, some sets of parameters may lead to better optimal decisions in learning under uncertainty.

3.2.8 The subjective experience-based decisions may be helpful or harmful in optimal learning under uncertainty

The subjective experience-based decisions may be helpful or harmful in optimal learning under uncertainty, depending on the underlying domain type and the frame employed by the decision maker. Simulation results on subjective discriminability are consistent with findings from human decision experiments in the psychology literature. Also, the results
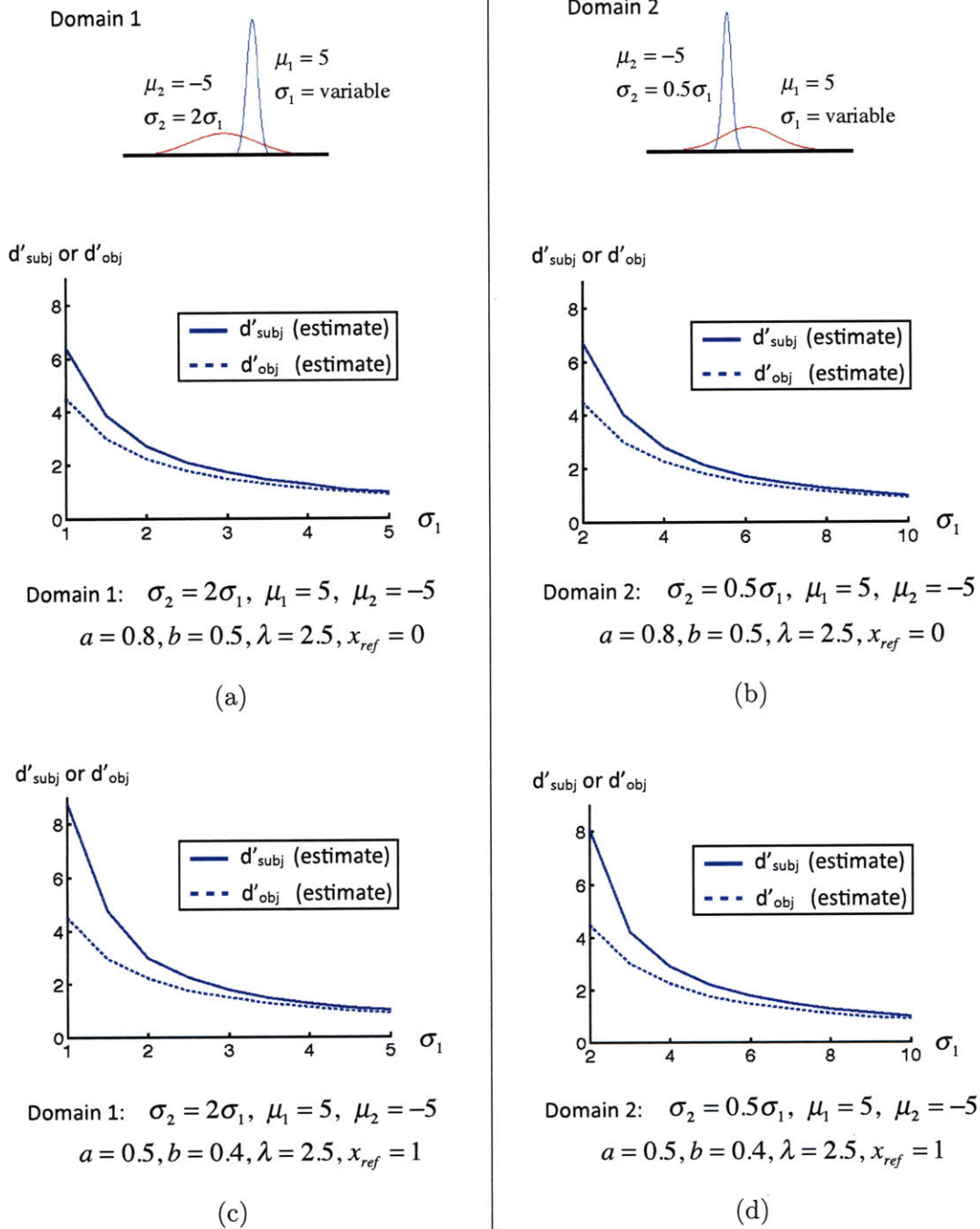
Figure 3-7: The influence of outcome variances on discriminabilities for each domain. Subplots (a) and (c): Domain 1 with $\sigma_2 = 2\sigma_1$. Subplots (b) and (d): Domain 2 with $\sigma_2 = 0.5\sigma_1$

were confirmed by the human decision experiments in Chapter 4.

On Domain 1 with the gain frame (Figure 3-3 (b)), the experience-based mode allows

the decision maker to detect the valence (good or bad) of each option quickly and correctly by increasing the subjective discriminability. Similarly, the human decision-making studies such as the IOWA gambling experiments have shown that humans with normal affective decision-making ability are good at quickly detecting the optimal decision in this kind of domain (where the optimal option involves low variance in outcomes and the suboptimal option involves high variance), whereas patients with the vmPFC damage are not [5, 4, 40]. Since most outcomes in IOWA experiments were gains (in particular, all outcomes from the optimal option were gains), normal participants would view the decision task in the gain frame. Thus, their quick optimal decisions on this domain employing the gain frame may be related to their subjective experience-based learning and the increased subjective discriminability. The loss-averse attitude and the risk-averse attitude in face of gains were helpful in optimal learning on this domain.

Subjective experience and emotional reactions, however, might impair decision making in the other domain. On Domain 2 with the gain frame (Figure 3-3 (d)), people often fail to detect the optimal option under uncertainty. Shiv *et al.*'s experiment [54] may be associated with this harmful side of the subjective experience-based learning in optimal decision behavior. In Shiv *et al.*'s experiment, the task involved 20 rounds of investment decisions between one safe option (no investment, i.e., $1 gain for sure) and the other risky option (investment, i.e., $3.5 gain with 50% chance and $1 loss with 50% chance). Interestingly, target patients (patients with stable focal lesions in brain regions related to emotion) made more advantageous decisions and ultimately earned more money from their investments than normal participants and control patients (patients with stable focal lesions in brain regions unrelated to emotion). Normal participants and control patients, more affected by the outcomes of decisions made in the previous rounds, adopted a conservative strategy and became more reluctant to invest on the subsequent round than target patients. This non-optimal decision-making behavior found from normal participants and control patients may arise from their risk-averse attitude in face of gains in their subjective experience-based learning.

3.3 Discriminability during exploitative trials

In sections 3.2.2 and 3.2.3, the concepts of objective and subjective discriminabilities with initial $2n_B$ exploratory trials were defined. In this section, the concept of discriminability after the initial exploratory trials is investigated regarding each of four different decision strategies (objective outcome-based greedy selection, probability matching (action value sampling), myopic VPI (value of perfect information), subjective value-based greedy selection).

3.3.1 Greedy selection based on objective outcomes

Suppose that after an initial $2n_B$ exploratory trials, the decision maker follows the greedy selection rule based on objective outcomes. Here it is investigated what concept of discriminability is associated with this greedy selection-based decisions.

The mean of sampled outcomes of option $k = 1$, 2 is denoted as $\hat{\mu}_k^t \triangleq (1/n_k^t) \sum_{i=1}^{n_k^t} r_k^{(i)}$ where n_k^t is the number of sampled outcomes of option k before trial t. Also, $(\hat{\sigma}_k^t)^2$ denotes the variance estimate of outcomes of option k on trial t: $(\hat{\sigma}_k^t)^2 = 1/(n_k^t - 1) \sum_{i=1}^{n_k^t} (r_k^{(i)} - \hat{\mu}_k^t)^2$.

On trial $t \geq t_B + 1$ the probability of choosing option 1 over option 2 can be computed as:

$$
\begin{aligned}
\mathrm{Pr}_{obj}^t(\text{option} = 1) &= \mathrm{Pr}_{obj}(a_t = 1) = \mathrm{Pr}(\hat{\mu}_1^t > \hat{\mu}_2^t) \\
&= \mathrm{Pr}(\hat{\mu}_1^t > \hat{\mu}_2^t | \hat{\mu}_1^{t-1} > \hat{\mu}_2^{t-1}) \mathrm{Pr}(\hat{\mu}_1^{t-1} > \hat{\mu}_2^{t-1}) + \mathrm{Pr}(\hat{\mu}_1^t > \hat{\mu}_2^t | \hat{\mu}_1^{t-1} < \hat{\mu}_2^{t-1}) \mathrm{Pr}(\hat{\mu}_1^{t-1} < \hat{\mu}_2^{t-1}) \\
&= \mathrm{Pr}(\hat{\mu}_1^t > \hat{\mu}_2^t | \hat{\mu}_1^{t-1} > \hat{\mu}_2^{t-1}) \mathrm{Pr}_{obj}(a_{t-1} = 1) + \mathrm{Pr}(\hat{\mu}_1^t > \hat{\mu}_2^t | \hat{\mu}_1^{t-1} < \hat{\mu}_2^{t-1})(1 - \mathrm{Pr}_{obj}(a_{t-1} = 1)) \\
&= \Phi(d_{G1}^t) \mathrm{Pr}_{obj}(a_{t-1} = 1) + \Phi(d_{G2}^t)(1 - \mathrm{Pr}_{obj}(a_{t-1} = 1))
\end{aligned}
$$

where $d_{G1}^t \simeq n_1^t \left(\frac{\hat{\mu}_1^{t-1} - \hat{\mu}_2^{t-1}}{\sigma_1} \right)$ and $d_{G2}^t \simeq n_2^t \left(\frac{\hat{\mu}_1^{t-1} - \hat{\mu}_2^{t-1}}{\sigma_2} \right)$.

The condition $\hat{\mu}_1^{t-1} > \hat{\mu}_2^{t-1}$ means that the decision maker selected option 1 on trial $t-1$ according to the greedy selection rule and obtained a new random outcome $r_1^{t-1} \sim \mathrm{N}(\mu_1, \sigma_1^2)$; thus, there is no change on the sample mean of option 2 (i.e., $\hat{\mu}_2^t = \hat{\mu}_2^{t-1}$). The sample mean of option 1, $\hat{\mu}_1^t$ is updated (given $\hat{\mu}_1^{t-1}$) as follows: $\hat{\mu}_1^t = \hat{\mu}_1^{t-1} + K_1^t(r_1^{t-1} - \hat{\mu}_1^{t-1})$ using the gain $K_1^t = 1/n_1^t$. Thus, $\hat{\mu}_1^t | \hat{\mu}_1^{t-1} \sim \mathrm{N}(\hat{\mu}_1^{t-1} + K_1^t(\mu_1 - \hat{\mu}_1^{t-1}), (K_1^t \sigma_1)^2)$. Taking it into account that μ_1^t is a normal random variable and μ_2^t is a fixed number given μ_1^{t-1}

and μ_2^{t-1} , the new random variable $y \triangleq \hat{\mu}_1^t - \hat{\mu}_2^t$ is also normal following $y \sim \mathrm{N}(\hat{\mu}_1^{t-1} - \hat{\mu}_2^{t-1} + K_1^t(\mu_1 - \hat{\mu}_1^{t-1}), (K_1^t \sigma_1)^2)$. Thus, $\Pr(\hat{\mu}_1^t > \hat{\mu}_2^t | \hat{\mu}_1^{t-1} > \hat{\mu}_2^{t-1}) = \Pr(y > 0) = \Phi(d_{G1}^t)$ where $d_{G1}^t = \frac{\hat{\mu}_1^{t-1} - \hat{\mu}_2^{t-1}}{\sigma_1 / n_1^t} + \frac{\mu_1 - \hat{\mu}_1^{t-1}}{\sigma_1} \simeq n_1^t \left(\frac{\hat{\mu}_1^{t-1} - \hat{\mu}_2^{t-1}}{\sigma_1} \right)$ (because $\frac{\hat{\mu}_1^{t-1} - \hat{\mu}_2^{t-1}}{\sigma_1 / n_1^t} >> \frac{\mu_1 - \hat{\mu}_1^{t-1}}{\sigma_1}$ for a sufficiently large n_1^t). In like manner, under the condition $\hat{\mu}_1^{t-1} < \hat{\mu}_2^{t-1}$, there is no change on the sample mean of option 1 (i.e., $\hat{\mu}_1^t = \hat{\mu}_1^{t-1}$). Also, $\hat{\mu}_2^t = \hat{\mu}_2^{t-1} + K_2^t(r_2^{t-1} - \hat{\mu}_2^{t-1})$ using the gain $K_2^t = 1/n_2^t$. Here, $\hat{\mu}_1^t$ is a fixed number and $\hat{\mu}_2^t$ is a random variable, that is, $\hat{\mu}_2^t | \hat{\mu}_2^{t-1} \sim \mathrm{N}(\hat{\mu}_2^{t-1} + K_2^t(\mu_2 - \hat{\mu}_2^{t-1}), (K_2^t \sigma_2)^2)$. The new random variable $y \triangleq \hat{\mu}_1^t - \hat{\mu}_2^t$ is also normal following $y \sim \mathrm{N}(\hat{\mu}_1^{t-1} - \hat{\mu}_2^{t-1} - K_2^t(\mu_2 - \hat{\mu}_2^{t-1}), (K_2^t \sigma_2)^2)$. Thus, the following relationships are obtained: $\Pr(\hat{\mu}_1^t > \hat{\mu}_2^t | \hat{\mu}_1^{t-1} < \hat{\mu}_2^{t-1}) = \Phi(d_{G2}^t)$ where $d_{G2}^t = \frac{\hat{\mu}_1^{t-1} - \hat{\mu}_2^{t-1}}{\sigma_2 / n_2^t} - \frac{\mu_2 - \hat{\mu}_2^{t-1}}{\sigma_2} \simeq n_2^t \left(\frac{\hat{\mu}_1^{t-1} - \hat{\mu}_2^{t-1}}{\sigma_2} \right)$ (because $\frac{\hat{\mu}_1^{t-1} - \hat{\mu}_2^{t-1}}{\sigma_2 / n_2^t} >> \frac{\mu_2 - \hat{\mu}_2^{t-1}}{\sigma_2}$ for a sufficiently large n_2^t).

3.3.2 Probability matching (action value sampling)

Suppose that after an initial $2n_B$ exploratory trials, the decision maker follows the probability matching (action value sampling) rule [17]. Here it is investigated what concept of discriminability is associated with this probability matching-based decisions.

The following update rule is applied to incorporate the observation r_k^{t-1} (the outcome of option k on trial $t-1$): When $t \geq t_B + 1$, $\hat{\mu}_k^t = \hat{\mu}_k^{t-1} + K_k^t(r_k^{t-1} - \hat{\mu}_k^{t-1})$ using the gain $K_k^t = 1/n_k^t$ where n_k^t is the number of observed outcomes of option k before trial t. Note that r_k^{t-1} is the n_k^tth observed outcome of option k, that is, $r_k^{(n_k^t)}$, and that $\hat{\mu}_k^{t_B}$ can be obtained from initial exploration. Also, when $t \geq t_B + 1$, $(\hat{s}_k^t)^2 = (1 - K_k^t)(\hat{s}_k^{t-1})^2$. Note that $(\hat{s}_k^{t_B})^2 = (\hat{\sigma}_k^{t_B})^2 / n_k^{t_B}$.

The probability distribution of outcome r_k for option k is normal: $r_k \sim \mathrm{N}(\mu_k, \sigma_k^2)$. Now the probability distribution of the mean μ_k on trial $t \geq t_B + 1$ given all past observed outcomes $r_k^{(i)} (i = 1, \cdots, n_k^t)$ can be described as $\mu_k^t \sim \mathrm{N}(\hat{\mu}_k^t, (\hat{s}_k^t)^2)$ where $(\hat{s}_k^t)^2 \simeq (\hat{\sigma}_k^t)^2 / n_k^t$.

On trial $t \geq t_B + 1$ the probability of choosing option 1 over option 2 can be computed as: $\Pr_{sampling}^t(\text{option} = 1) = \Pr(\mu_1^t > \mu_2^t) = \Pr(\mu_1^t - \mu_2^t > 0) = \Pr(y > 0)$ where $y \triangleq \mu_1^t - \mu_2^t$. Thus, $\Pr_{sampling}^t(\text{option} = 1) = \Phi(d^t)$ where $d^t = \frac{\hat{\mu}_1^t - \hat{\mu}_2^t}{\sqrt{(\hat{\sigma}_1^t)^2 / n_1^t + (\hat{\sigma}_2^t)^2 / n_2^t}}$.

3.3.3 Myopic VPI (Value of Perfect Information) selection

Suppose that after an initial $2n_B$ exploratory trials, the decision maker follows the greedy selection (based on objective outcomes) *with myopic VPI* (value of perfect information) [17]. Here it is investigated what concept of discriminability is associated with this myopic VPI-based decisions.

The probability distribution of the mean μ_k on trial t given all past observed outcomes $r_k^{(i)}(i = 1, \cdots, n_k^t)$ can be described as $\mu_k^t \sim N(\hat{\mu}_k^t, (\hat{s}_k^t)^2)$ where $(\hat{s}_k^t)^2 \simeq (\hat{\sigma}_k^t)^2/n_k^t$.

For option pair $(u, v) = (1, 2)$ or $(2,1)$, defining $x \overset{\Delta}{=} \mu_u^t$, then $p(x; \hat{\mu}_u^t, (\hat{s}_u^t)^2) = N(\hat{\mu}_u^t, (\hat{s}_u^t)^2)$. When $\hat{\mu}_u^t > \hat{\mu}_v^t$, $VPI_u^t = \int_{-\infty}^{\hat{\mu}_v^t}(\hat{\mu}_v^t - x)\, p(x; \hat{\mu}_u^t, (\hat{s}_u^t)^2)\ dx = -(\hat{\mu}_u^t - \hat{\mu}_v^t)\, \Phi(-\frac{\hat{\mu}_u^t - \hat{\mu}_v^t}{\hat{s}_u^t}) + \hat{s}_u^t\, \phi(\frac{\hat{\mu}_u^t - \hat{\mu}_v^t}{\hat{s}_u^t})$. When $\hat{\mu}_u^t < \hat{\mu}_v^t$, $VPI_u^t = \int_{\hat{\mu}_u^t}^{\infty}(x - \hat{\mu}_u^t)\, p(x; \hat{\mu}_u^t, (\hat{s}_u^t)^2)\ dx = (\hat{\mu}_u^t - \hat{\mu}_v^t)\, \Phi(\frac{\hat{\mu}_u^t - \hat{\mu}_v^t}{\hat{s}_u^t}) + \hat{s}_u^t\, \phi(\frac{\hat{\mu}_u^t - \hat{\mu}_v^t}{\hat{s}_u^t})$. Thus, $VPI_u^t = -|\hat{\mu}_u^t - \hat{\mu}_v^t|)\, \Phi(-\frac{|\hat{\mu}_u^t - \hat{\mu}_v^t|}{\hat{s}_u^t}) + \hat{s}_u^t\, \phi(\frac{\hat{\mu}_u^t - \hat{\mu}_v^t}{\hat{s}_u^t})$ for any option pair (u, v). Note that VPI can be viewed as a sort of exploration bonus provided to outcome uncertainty. That is, the option with a greater outcome uncertainty will be selected more often due to the exploration bonus. The idea of uncertainty-based exploration strategy arises from a Bayesian formulation in which the new information gathered from the option with greater outcome uncertainty is more likely to change the future decision strategy, compared to that from options with smaller outcome uncertainty.

Using a similar method as in Section 3.3.1, on trial $t \geq t_B + 1$ the probability of choosing option 1 over option 2 can be computed as:

$$\Pr_{vpi}^t(\text{option} = 1) = \Pr_{vpi}(a_t = 1) = \Pr(\hat{\mu}_1^t + VPI_1^t > \hat{\mu}_2^t + VPI_2^t)$$

$$= \Phi(d_{G1}^t)\, \Pr_{vpi}(a_{t-1} = 1) + \Phi(d_{G2}^t)(1 - \Pr_{vpi}(a_{t-1} = 1))$$

where $d_{G1}^t \simeq n_1^t\left(\frac{\hat{\mu}_1^{t-1} - \hat{\mu}_2^{t-1} + (VPI_1^t - VPI_2^t)}{\sigma_1}\right)$ and $d_{G2}^t \simeq n_2^t\left(\frac{\hat{\mu}_1^{t-1} - \hat{\mu}_2^{t-1} + (VPI_1^t - VPI_2^t)}{\sigma_2}\right)$.

3.3.4 Greedy selection based on subjective values

This approach is exactly the same as in Section 3.3.1, except for employing subjective values instead of objective outcomes.

The mean of sampled subjective values of option k ($= 1, 2$) is denoted as $\hat{\mu}_{subj,k}^t \overset{\Delta}{=} (1/n_k^t)\sum_{i=1}^{n_k^t} v_k^{(i)}$ where $v_k^{(i)} = f(r_k^{(i)})$ and n_k^t is the number of sampled outcomes of option k ($= 1, 2$) before trial t. Also, $(\hat{\sigma}_{subj,k}^t)^2$ denotes the variance estimate of subjective values of option k on trial t: $(\hat{\sigma}_{subj,k}^t)^2 = 1/(n_k^t - 1)\sum_{i=1}^{n_k^t}(v_k^{(i)} - \hat{\mu}_{subj,k}^t)^2$.

85

On trial $t(\geq t_B + 1)$ the probability of choosing option 1 over option 2 can be computed as:

$$\Pr^t_{subj}(\text{option} = 1) = \Pr_{subj}(a_t = 1) = \Pr(\hat{\mu}^t_1 > \hat{\mu}^t_2)$$

$$= \Pr(\hat{\mu}^t_1 > \hat{\mu}^t_2 | \hat{\mu}^{t-1}_1 > \hat{\mu}^{t-1}_2) \Pr(\hat{\mu}^{t-1}_1 > \hat{\mu}^{t-1}_2) + \Pr(\hat{\mu}^t_1 > \hat{\mu}^t_2 | \hat{\mu}^{t-1}_1 < \hat{\mu}^{t-1}_2) \Pr(\hat{\mu}^{t-1}_1 < \hat{\mu}^{t-1}_2)$$

$$= \Pr(\hat{\mu}^t_1 > \hat{\mu}^t_2 | \hat{\mu}^{t-1}_1 > \hat{\mu}^{t-1}_2) \Pr_{subj}(a_{t-1} = 1) + \Pr(\hat{\mu}^t_1 > \hat{\mu}^t_2 | \hat{\mu}^{t-1}_1 < \hat{\mu}^{t-1}_2)(1 - \Pr_{subj}(a_{t-1} = 1))$$

$$= \Phi(d^t_{G1}) \Pr_{subj}(a_{t-1} = 1) + \Phi(d^t_{G2})(1 - \Pr_{subj}(a_{t-1} = 1))$$

where $d^t_{G1} \simeq n^t_1 \left(\frac{\hat{\mu}^{t-1}_1 - \hat{\mu}^{t-1}_2}{\sigma_1} \right)$ and $d^t_{G2} \simeq n^t_2 \left(\frac{\hat{\mu}^{t-1}_1 - \hat{\mu}^{t-1}_2}{\sigma_2} \right)$.

Figure 3-8 shows the probability of selecting the optimal option over trials on each domain (Domain 1 and Domain 2) for different decision strategies (Subj: subjective value-based greedy selection, Obj: objective outcome-based greedy selection, Sampling: probability matching, VPI: myopic VPI). In both simulations (Domain 1 and Domain 2), each strategy had an initial 10 exploratory trials ($n_B = 5$ trials for each option). Also, the reference points were set to the mean outcome of the risky option (i.e., the gain frame on Domain 1 and the loss frame on Domain 2) to obtain an increased subjective discriminability, according to the findings in Section 3.2.6. Simulation results showed that on both domains the subjective value-based greedy selection strategy with those reference point selections led to the greatest probability of selecting the optimal option over trials (and thus, the greatest total outcome), compared to other strategies.

3.4 Affective reinforcement learning for Markov decision processes (MDPs)

Subjective and affective elements are well-known to influence human learning and decision making. The research for exploring and exploiting these important influences in computational learning theory, however, is still in its early stage. This section presents a new model combining subjective and affective influences within the RL and MDP framework. The affective-congitive (AC) model involves two different modes: the experience-based mode and the prediction-based mode. To model the total-experienced utility from past experiences, the AC model introduces a prospect theory (PT)-based parameterized "experienced-utility
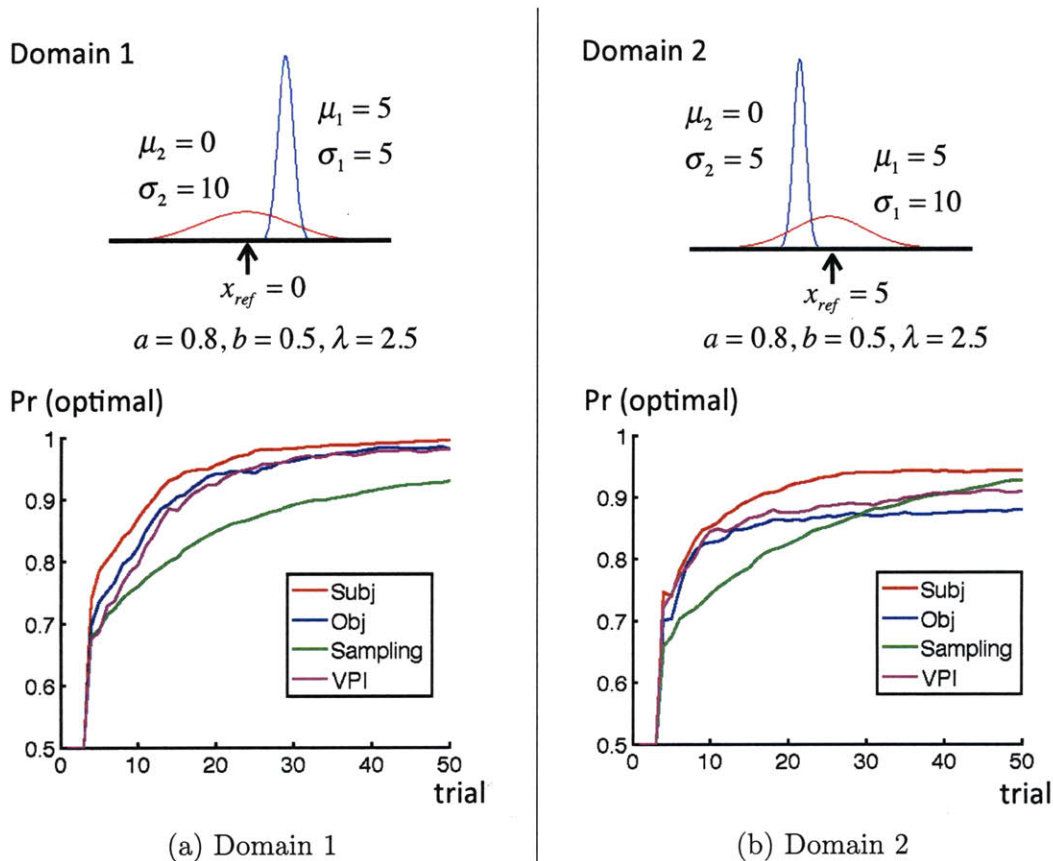
Figure 3-8: The probability of selecting the optimal option over trials on each domain for different decision strategies (Subj: subjective value-based greedy selection, Obj: objective outcome-based greedy selection, Sampling: probability matching, VPI: myopic VPI)

function". In order to model affective-subjective characteristics of prediction-based mode, the AC model employs a prospect theory (PT)-based parameterized "predicted-utility function". It also models one specific kind of affective state, called the "goal-achieving (confidence) state," which relates to the sense of confidence in the current decision-making policy. In economics theory, the PT value function is fixed, but it is hypothesized that the affective state influences the shape of the predicted-utility function (i.e., sensitivities to the expected gains and losses). An RL-based computational framework that implements this hypothesis automatically regulates trade-offs between exploration and exploitation while beating the performance of five other well-known model-free learning algorithms.

The AC model includes both a subjective component (PT value function) and a compo-

nent that captures part of how affective states may influence decision making. It is further hypothesized that the latter component can influence the former to give performance that is closer to human behavior. Furthermore, while PT theory has been developed in economics for domains with known outcome distributions, this new model enables PT theory to be used for unknown and changing stochastic outcome distributions. Finally, it is known that in the face of multiple unknown nonstationary distributions of outcomes, balancing the trade-offs between exploration and exploitation is very critical; this new model achieves this balanced trade-off in an automatic and internally-regulated way.

3.4.1 Markov Decision Processes

A discrete-time finite MDP is a tuple $M = (S, A, T, \gamma, R)$, where S is a finite set of states, A is a finite set of actions, $p_{ss'}^a = \Pr\{s_{t+1} = s' | s_t = s, a_t = a\}$ are one-step state-transition probabilities when taking action a in state s, and $\gamma \in [0, 1)$ is the discount factor, and $r_s^a = E\{r_{t+1} | s_t = s, a_t = a\}$ are one-step expected outcomes [58]. A stochastic policy is a mapping from states to probabilities of taking each action $\pi : S \times A \mapsto [0, 1]$. The *state-value function* $V^\pi(s)$ is the expected discounted future outcome from each state s:

$$
\begin{aligned}
V^\pi(s) &= E_\pi[r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \cdots | s_t = s] \\
&= E_\pi[r_{t+1} + \gamma V^\pi(s_{t+1}) | s_t = s] \\
&= \sum_a \pi(s, a)[r_s^a + \gamma \sum_{s'} p_{ss'}^a V^\pi(s')]
\end{aligned}
\tag{3.1}
$$

The *action-value function* $Q^\pi(s, a)$ is the expected discounted future outcome for taking a in s and thereafter following π:

$$
\begin{aligned}
Q^\pi(s, a) &= E_\pi[r_{t+1} + \gamma r_{t+2} + \cdots | s_t = s, a_t = a] \\
&= E_\pi[r_{t+1} + \gamma V^\pi(s_{t+1}) | s_t = s, a_t = a] \\
&= r_s^a + \gamma \sum_{s'} p_{ss'}^a V^\pi(s')
\end{aligned}
\tag{3.2}
$$

Note that $V^\pi(s) = \sum_a \pi(s,a) Q^\pi(s,a)$. The optimal action-value function is

$$Q^*(s,a) = \max_\pi Q^\pi(s,a) = r_s^a + \gamma \sum_{s'} p_{ss'}^a \max_{a'} Q^*(s',a').$$

3.4.2 Affective Reinforcement Learning

Suppose that the agent with a policy π takes action $a_t \in A$ at state $s_t \in S$ in timestep t, and then reaches next state $s_{t+1} \in S$ and obtains the immediate outcome r_{t+1}. For this experience tuple $< s_t, a_t, r_{t+1}, s_{t+1} >$, the estimated action-value function Q uses the following Q-learning based update rule to approximate the optimal action-value function Q^*:

$$
\begin{aligned}
R_t &= r_{t+1} + \gamma \ \max_{a'} Q(s_{t+1}, a') \\
Q(s_t, a_t) &\leftarrow Q(s_t, a_t) + \alpha \ (R_t - Q(s_t, a_t))
\end{aligned}
\tag{3.3}
$$

where α is a learning rate parameter. Thus, $Q(s,a)$ estimates the expected long-term future outcome for taking action a at state s. Note that R_t is a sampled estimate for the long-term future outcome. In the proposed model, I call this sort of action-value estimation the "deliberative estimation" phase.

Deliberative estimation estimates the expectation of the long-term future (objective) outcome for each candidate action. Although I present a model-free Q-learning method for this estimation here, any model-based learning methods can be employed if they are useful for approximating the long-term outcomes.

In contrast to traditional RL algorithms in which the expected future outcome is directly used for the action-selection model (decision-making policy), the proposed algorithm computes predicted utility, which models affective-subjective prediction in decision making. This additional step, that is, transforming an expected future outcome into a predicted utility, is called the "affective shaping" phase.

Affective shaping approximates the influence of the current affective state, at the moment of prediction and decision making, on the predicted utility. Thus, in the proposed model, computing the predicted utility for each candidate action is approximated by two phases: the "deliberative estimation" phase and the "affective shaping" phase. I model the "goal-achieving (confidence) state $e(s)$," which is a kind of task-related affective state. This

(a) High Confidence State   (b) Low Confidence State
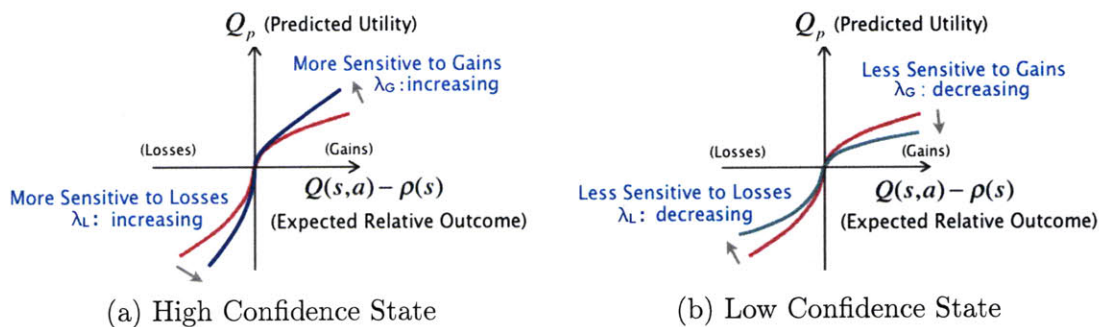
Figure 3-9: The shape of the predicted-utility function changes with the confidence state.

confidence state is analogous to how the agent "feels" about the current policy in terms of achieving the goal (in this case, how confident (or not) the agent is using the current policy).

In learning situations the agent should regulate the trade-offs between exploration and exploitation [58, 13]. To model this exploratory regulation, I focus on a policy based on the Boltzmann distribution. The agent selects action a at state s with the probability:

$$\pi(s,a) = \frac{\exp(\beta\,Q_D(s,a))}{\sum_{i \in A} \exp(\beta\,Q_D(s,i))} \tag{3.4}$$

where β is called the *inverse temperature* parameter.

$Q_D(s,a)$ is called *decision utility*, which the agent actually employs to compute the policy $\pi(s,a)$. Decision utility can be thought of as the preference for taking action a in state s. Note that as β is greater, the selection of actions with higher Q_D is more likely. Since β globally influences the level of exploration at all states, β can be viewed as a "global" control parameter for the balance between exploration and exploitation. In contrast, the confidence state $e(s)$ is a "local" control parameter, which can be viewed as the *local inverse temperature* parameter. The following four assumptions are made for the affective RL framework, summarized in Figure 3-10.

Assumption 1. The agent's decision utility $Q_D(s,a)$ is the linear combination of experienced utility $Q_E(s,a)$ and predicted utility $Q_P(s,a)$: $Q_D(s,a) = (1-\eta)Q_E(s,a) + \eta\,Q_P(s,a)$

90

Parameters: β (inverse temperature), α (learning rate), κ (emotion sensitivity), η (cognitive control), Q_0 (initial action values), γ (discount factor).

Initialization: action values $Q(s,a) = Q_0$ and reference points $\rho(s) = Q_0$ for all state s and action a.

(1): For current state s, choose action a according to policy $\pi(s,a)$ in Equation 3.4.

(2): Take action a, and then reach new state s' and obtain outcome r.

(3): Update action value $Q(s,a)$ according to the deliberative estimation method, such as in Equation 3.3.

(4): Update action values on history paths up to state s according to Equation 3.10.

(5): Update total-experienced utility $Q_E(s,a)$ according to Equation 3.8.

(6): Update reference point $\rho(s)$ according to the reference point model, such as in Equation 3.6.

(7): Update confidence state $e(s)$ according to Equation 3.9.

(8): Compute predicted utility $Q_P(s,a)$ according to Equation 3.5.

(9): Compute decision utility $Q_D(s,a) = (1 - \eta)Q_E(s,a) + \eta\, Q_P(s,a)$.

(10): Go to (1).

Figure 3-10: The Affective RL Framework

where $\eta \in [0,1]$ is a cognitive control parameter between the experience-based mode and the prediction-based mode. For example, when $\eta = 0$ (e.g., the computational (cognitive) load is extremely high), the agent can rely only on the experienced-based mode (i.e., $Q_D = Q_E$).

Assumption 2. The predicted utility $Q_P(s,a)$ can be modeled by the predicted-utility function, a parameterized PT value function to incorporate affective-subjective influences into the agent's prediction. I focus on the *confidence state* $e(s)$, which is relevant for exploratory regulation. Thus, the predicted utility $Q_P(s,a)$ can be computed by the predicted-utility function $f_P(\cdot\,|e(s))$ whose shape (i.e., sensitivities to losses and gains) depends on the confidence state $e(s) \in [0,1]$, as follows: $Q_P(s,a) = f_P(Q(s,a) - \rho(s)|e(s))$ where $Q(s,a)$ is the estimated action value and $\rho(s)$ is the reference point (explained below). For $v = Q(s,a) - \rho(s)$, the expected relative (long-term) outcome,

$$f_P(v \mid e(s)) = \begin{cases} -\lambda_G(e(s))\, v^a, & v \geq 0 \\ -\lambda_L(e(s))\, (-v)^b, & v < 0 \end{cases} \tag{3.5}$$

91

$$\lambda_G(e(s)) = \lambda_{Gbase} + \lambda_{Gslope}(2\,e(s) - 1)$$

$$\lambda_L(e(s)) = \lambda_{Lbase} + \lambda_{Lslope}(2\,e(s) - 1)$$

where λ_{Gbase}, λ_{Gslope}, λ_{Lbase}, λ_{Lslope} are the parameters that determine how the sensitivities of the predicted-utility function change with $e(s)$.

The following two hypotheses are assumed:

(1) As the agent becomes more confident in achieving the goal (i.e., higher $e(s)$), it becomes more sensitive to potential gains and losses. See Figure 3-9 (a).

(2) As the agent becomes less confident in achieving the goal (i.e., lower $e(s)$), it becomes less sensitive to potential gains and losses. See Figure 3-9 (b).

Thus, $a > 0$, $b > 0$, $\lambda_{Gbase} > \lambda_{Gslope} > 0$ and $\lambda_{Lbase} > \lambda_{Lslope} > 0$. Note that the function is a loss-aversive piecewise linear function if $a = 1$, $b = 1$, $\lambda_{Gbase} = 1$ and $\lambda_{Lbase} > 1$. Experiments below will show that the first hypothesis will tend to lead to well-timed choices of using the strategy of exploitation, while the second hypothesis will tend to lead to well-timed choices of using the strategy of exploration.

Assumption 3. The reference point $\rho(s)$ is modeled as the average of the currently-estimated best action $a_{\text{first}}(s) = \arg\max_{i \in A} Q(s, i)$ and second best action $a_{\text{second}}(s) = \arg\max_{\substack{i \in A, \\ i \neq a_{\text{first}}}} Q(s, i)$ at state s:

$$\rho(s) = (Q(s, a_{\text{first}}(s)) + Q(s, a_{\text{second}}(s)))/2 \tag{3.6}$$

This model appears to have a nice algorithmic advantage of making the predicted utility of $a_{\text{first}}(s)$ positive and those of all other choices negative.

Assumption 4. The total-experienced utility $Q_E(s, a)$ can be modeled by the average of sampled (long-run) experienced utilities: $Q_E(s, a) = E_\pi[f_E(R_t - \rho(s_t))|s_t = s, a_t = a]$ where f_E is the experienced-utility function, another PT-based subjective value function that is different from the predicted-utility function f_P. Here it is assumed that f_E is the same as f_P but with a fixed neutral confidence ($e(s) = 0.5$); thus, f_E is a fixed-shape function independent of $e(s)$.

Experienced utility X_t from a sampled long-term relative outcome $v_t = R_t - \rho(s_t)$ is

92

computed as follows:

$$X_t = f_E(v_t) = \begin{cases} -\lambda_{Gbase} \; v_t^a, & v_t \geq 0 \\ -\lambda_{Lbase} \; (-v_t)^b, & v_t < 0 \end{cases} \tag{3.7}$$

where λ_{Gbase}, λ_{Lbase} are the parameters that determine the sensitivities of the experienced-utility functions.

For the experience tuple $< s_t, a_t, r_{t+1}, s_{t+1} >$, total-experienced utility $Q_E(s_t, a_t)$ is updated as follows:

$$Q_E(s_t, a_t) \leftarrow Q_E(s_t, a_t) + \alpha \; (X_t - Q_E(s_t, a_t)) \tag{3.8}$$

where α is a learning rate parameter.

Assumption 5. The confidence state $e(s)$ can be modeled by the following update rule: for an experience tuple $< s_t, a_t, r_{t+1}, s_{t+1} >$,

$$R_t = r_{t+1} + \gamma \; \max_{a'} Q(s_{t+1}, a')$$

$$\text{if } a_t = a_{\text{first}}(s_t), \quad q(s_t) \leftarrow q(s_t) + X_t \tag{3.9}$$

$$\text{otherwise}, \quad q(s_t) \leftarrow q(s_t) - X_t$$

$$e(s_t) \leftarrow \frac{1}{1 + exp(-\kappa q(s_t))}$$

where $\kappa > 0$ is a sensitivity parameter.

Note that $0 \leq e(s) \leq 1$ where 0 is analogous to the agent "feeling bad" about the current policy, 0.5 is neutral, and as $e(s)$ gets closer to 1, the value represents the agent "feeling better," being more confident the current policy will achieve the desired goal.

Action-value updates based on history paths. For each state s, the agent keeps N_{path} kinds of history paths of length L_{path} up to state s in memory for later action-value updates whenever the agent reaches state s again. For example, one history path of experience tuples up to state $s(= s_t)$ is $\{< s_{t-L}, a_{t-L}, r_{t-L+1}, s_{t-L+1} >, \cdots, < s_{t-2}, a_{t-2}, r_{t-1}, s_{t-1} >, < s_{t-1}, a_{t-1}, r_t, s_t >\} \equiv \{< s_{(1)}, a_{(1)}, r_{(2)}, s_{(2)} >, \cdots, < s_{(L-1)}, a_{(L-1)}, r_{(L)}, s_{(L)} >, < s_{(L)}, a_{(L)}, r_{(L+1)}, s_{(L+1)} >\}$
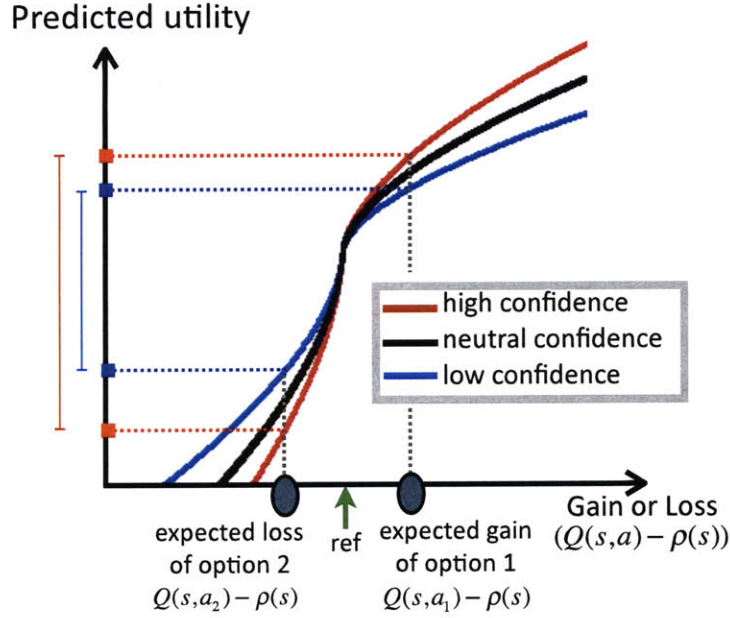
Figure 3-11: The Influence of Affective Shaping on Prediction: for instance, the case of two actions (two sky dots show the expected relative outcomes of two actions.)

At state $s_t \equiv s_{(L+1)}$, the agent updates action values on each history path:

For $k = L, \cdots, 1,$

$$R_{(k)} = r_{(k+1)} + \gamma \, max_{a'} Q(s_{(k+1)}, a') \tag{3.10}$$

$$Q(s_{(k)}, a_{(k)}) \leftarrow Q(s_{(k)}, a_{(k)}) + \alpha \, (R_{(k)} - Q(s_{(k)}, a_{(k)}))$$

3.4.3 Balancing the Trade-offs between Exploration and Exploitation

The confidence state $e(s)$ is locally defined for each state s. In Equation 3.9, if $X_t > 0$ (pleasurable experience) for the exploitative action ($a_t = a_{\text{first}}(s_t)$), or if $X_t < 0$ (displeasurable experience) for exploratory actions ($a_t \neq a_{\text{first}}(s_t)$), then $q(s_t)$ increases and thus confidence $e(s_t)$ increases; that is, the agent becomes more confident in the current *local* policy at s_t and will be more likely to take the exploitative action at that state in the future trials. However, if $X_t < 0$ for the exploitative action or $X_t > 0$ for exploratory actions, $q(s_t)$ decreases and thus confidence $e(s_t)$ decreases; that is, the agent becomes less confident in the current *local* policy at s_t and will be more likely to take exploratory actions at that

state in future trials.

Note that the sensitivity changes of the predicted-utility function according to $e(s)$ (Equation 3.5) help automatically balance the trade-offs between exploration and exploitation at state s. When the agent is more confident in the current policy ($e(s) > 0.5$, high confidence state), the predicted-utility function becomes more sensitive to expected gains and losses. Because of this larger sensitivity in a confident state, there are larger separations in the predicted utilities of the exploitative action and other actions. That is, compared to a neutral confidence state ($e(s) = 0.5$) or a low confidence state ($e(s) < 0.5$), the agent in a high confidence state predicts as if the estimated values of the exploitative action and other actions were more separated (See Figure 3-11); thus, the agent favors exploitation. When the agent becomes less confident, however, it has a less sensitive predicted-utility function in both domains. Because of this smaller sensitivity, there are smaller separations in the predicted utilities of the exploitative action and the other actions; thus, it promotes exploratory actions. This is made formal in the following theorem:

Theorem 1 Assume that there are K available actions $a_k \in A = \{a_1, \cdots, a_K\}$ at state s satisfying $Q(s, a_k) \geq Q(s, a_{k+1})$ for $k = 1, \cdots, K - 1$. Let the reference point $\rho(s)$ be given by $\rho(s) = (Q(s, a_1)) + Q(s, a_2))/2$. Then, the exploitation-exploration ratio at state s denoted as $w(\pi(s)) \equiv \frac{\pi(s, a_1)}{\sum\limits_{k=2}^{K} \pi(s, a_k)}$ increases with higher $e(s)$, and decreases with lower $e(s)$.

(Proof) The exploitation-exploration ratio at state s, $w(\pi(s))$ is:

$$w(\pi(s)) \equiv \frac{\pi(s, a_1)}{\sum\limits_{k=2}^{K} \pi(s, a_k)} = \frac{\exp(\beta Q_P(s, a_1))}{\sum\limits_{k=2}^{K} \exp(\beta Q_P(s, a_k))}$$

$$= \exp\left(\beta(Q_P(s, a_1) - Z)\right).$$

Let $\Delta = (Q(s, a_1) - Q(s, a_2))/2$. Note $\Delta \geq 0$. Assuming the reference point $\rho(s) =$

$(Q(s, a_1)) + Q(s, a_2))/2,$

$$Q(s, a_1) - \rho(s) = (Q(s, a_1) - Q(s, a_2))/2 = \Delta$$

$$Q(s, a_2) - \rho(s) = -(Q(s, a_1) - Q(s, a_2))/2 = -\Delta$$

From Equation 3.5, the predicted utilities are:

$$Q_P(s, a_1) = \lambda_G(e(s))(Q(s, a_1) - \rho(s))^a = \lambda_G(e(s))\Delta^a$$

$$Q_P(s, a_2) = -\lambda_L(e(s))(-(Q(s, a_2) - \rho(s)))^b = -\lambda_L(e(s))\Delta^b$$

Also, since $Q_P(s, a_k) \leq Q_P(s, a_2)$ for each $k(= 3, \cdots, K)$, $Q_P(s, a_k) = -Z_k + Q_P(s, a_2)$ using a non-negative constant Z_k (≥ 0) for each k.

$$Q_P(s, a_k) = -\lambda_L(e(s))(-(Q(s, a_k) - \rho(s)))^b = -\lambda_L(e(s))(-(-Z_k + Q(s, a_2) - \rho(s)))^b$$

$$= -\lambda_L(e(s))(Z_k + \Delta)^b$$

Note that $\lambda_G(e(s))$ and $\lambda_L(e(s))$ in Equation 3.5 increase with higher $e(s)$. Thus, $Q_P(s, a_1)$ increases with higher $e(s)$, and $Q_P(s, a_k)(k = 2, \cdots, K)$ decrease with higher $e(s)$. This means that the numerator of $w(\pi(s))$, $\exp(\beta Q_P(s, a_1))$ increases with higher $e(s)$, whereas the denominator of $w(\pi(s))$, $\sum_{k=2}^{K} \exp(\beta Q_P(s, a_k))$ decreases with higher $e(s)$. That is, the exploitation-exploration ratio $w(\pi(s))$ increases with higher $e(s)$ at state s. $\quad$ Q.E.D.

The prediction-based mode interacts with the experience-based mode in agent's learning and decision making. The affective RL model assumes that the agent's decision utility arises from the feedback of past subjective experiences in similar choice situations (total-experienced utility $Q_E(s, a)$, "experience-based mode") and the affective prediction about the future experienced utility of choice outcomes (predicted utility $Q_P(s, a)$, "prediction-based mode"): $Q_D(s, a) = (1 - \eta)Q_E(s, a) + \eta Q_P(s, a)$ where $\eta \in [0, 1]$ is a model parameter related to computational (cognitive) load. Psychological experiments on human decision making have shown that trade-offs between the two modes may depend on cognitive load

at the time of decision making as well as personal disposition [53]. That is, when people are stressful or more cognitively loaded, they tend to rely more on the experience-based mode $(\eta = 0)$.

The affective regulation using $e(s)$ is very useful for regulating the balance between exploration and exploitation in unpredictable dynamic worlds. When the world is temporarily changing, the agent becomes less confident in its current exploitative action and $e(s)$ facilitates exploration ("*affective heating*"). When the world is temporarily stationary, however, the agent becomes more confident in its current exploitative action and $e(s)$ facilitates exploitation ("*affective cooling*").

While simulated annealing typically moves only from hot to cold, the new affective regulation dynamically adjusts the local inverse temperature $e(s)$ to regulate the annealing, inducing momentary increases in temperature (analogous to "feeling uneasy") to favor more exploration in a changing world. If the agent finds, for example, that it is stuck in a local minimum, then this internal mechanism "heats up" the process automatically, allowing it to once again explore outside that minimum. Thus, this model of affect allows the learning system to dynamically regulate its behavior and improve its performance as it learns.

3.4.4 EXPERIMENTS

Several well-known RL algorithms will be compared with the affective RL algorithm in terms of the learning performance in three problems (from [17]) devised for testing exploratory regulation: (1) Q-learning with semi-uniform random exploration (2) Q-learning with Boltzmann exploration (3) Q-learning with Interval Estimation [27] (4) IEQL+ [39] (5) Bayesian Q-learning with VPI + mixture updating [17] (6) the affective RL model with only the experience-based mode (i.e., $\eta = 0$) (7) the affective RL model with only the prediction-based mode (i.e., $\eta = 1$) (8) the affective RL model with both experienced-based and prediction-based modes. I test in the following domains: **Chain, Loop** and **Maze**.

Chain: Figure 3-12 (a) shows the "Chain" problem that has five states and two actions a and b. With probability 0.2, the agent "slips" and its action actually has the consequence of the other action. The optimal policy of this problem is to choose action a at all states, even though the agent sometimes slips. In order to maximize the total return for a given

97

(a) Chain

(b) Loop

(c) Maze
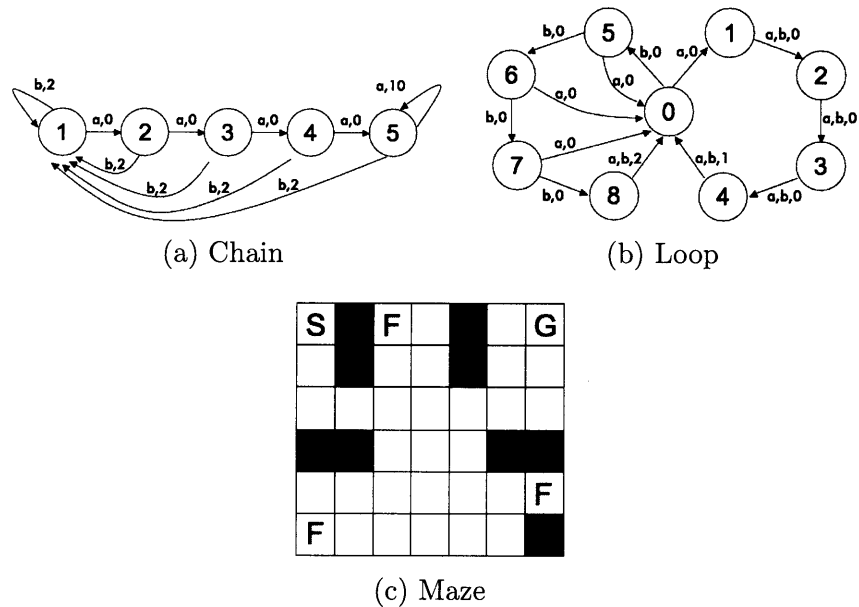
Figure 3-12: Three Test Problems

number of trials, the agent should not be trapped at the start state (state 1). Thus, the agent needs to balance the trade-offs between exploration and exploitation effectively.

Loop: Figure 3-12 (b) shows the "Loop" problem that has two loops linked at the start state (state 0). There are two deterministic actions a and b. The optimal policy of this problem is to choose action b at all states. The inferior exploration policies will converge too fast on action a for state 0.

Maze: Figure 3-12 (c) shows the "Maze" problem where the agent's goal is to collect flags and reach the goal. The start state is marked as S, the goal state is G, and each flag is located in one of three F states. When it gets to the goal state G, the agent obtains the reward based on the number of flags collected, and then is immediately moved to the start state. The agent can move in one of four directions: right, left, up and down. Also, with probability 0.1, the agent "slips" and actually moves in a perpendicular direction. If it attempts to move into a wall, it does not move. The maze has 33 discrete positions and there are eight kinds of combinations that represent the status of the three flags at any moment; thus, we assume that this problem has a total of 264 states.

In affective RL-based implementations, the following parameters were used for utility

CHAIN	1st Phase	2nd Phase
Uniform	1519.0 ± 37.2	1611.4 ± 34.7
Boltzmann	1605.8 ± 78.1	1623.4 ± 67.1
Interval	1522.8 ± 180.2	1542.6 ± 197.5
IEQL+	2343.6 ± 234.4	2557.4 ± 271.3
Bayes VPI+MIX	817.6 ± 101.8	1099.5 ±134.9
Affective RL (experience-based)	3381.3 ± 324.9	3576.6 ± 354.7
Affective RL (prediction-based)	3321.5 ± 317.6	3604.3 ± 311.1
Affective RL (both modes)	3395.9 ± 333.9	3708.9 ± 261.1

LOOP	1st Phase	2nd Phase
Uniform	185.6 ± 3.7	198.3 ± 1.4
Boltzmann	186.0 ± 2.8	200.0 ± 0.0
Interval	198.1 ± 1.4	200.0 ± 0.0
IEQL+	264.3 ± 1.6	292.8 ± 1.3
Bayes VPI+MIX	326.4 ± 85.2	340.0 ± 91.7
Affective RL (experience-based)	387.7 ± 6.3	400 ± 0.0
Affective RL (prediction-based)	391.0 ± 2.5	400 ± 0.0
Affective RL (both modes)	391.2 ± 2.5	400.0 ± 0.0

MAZE	1st Phase	2nd Phase
Uniform	105.3 ± 10.3	161.2 ± 8.6
Boltzmann	195.2 ± 61.4	1024.3 ± 87.9
Interval	246.0 ± 122.5	506.1 ± 315.1
IEQL+	269.4 ± 3.0	253.1 ± 7.3
Bayes VPI+MIX	817.6 ± 101.8	1099.5 ±134.9
Affective RL (experience-based)	1332.1 ± 103.3	1596.7 ± 27.7
Affective RL (prediction-based)	1543.2 ± 140.6	1654.9 ± 133.0
Affective RL (both modes)	1554.1 ± 167.4	1772.0 ± 102.9

Table 3.1: The Average and Standard Deviation of Accumulated Rewards for Each Phase

functions: $a = 0.8$, $b = 0.5$, $\lambda_{Gbase} = 1$, $\lambda_{Gslope} = 1$, $\lambda_{Lbase} = 2.5$, $\lambda_{Lslope} = 1$ [1] (in Equations 3.5 and 3.7). The discount factor $\gamma = 0.99$ was used. Also, two history paths of length 5 for each state were kept in memory for action value updates.

Although a fixed inverse temperature β is used without any global annealing schedule for it, the agent could automatically regulate the trade-offs between exploration and exploitation through the local goal-achieving affective state $e(s)$. To measure the learning performance, parameters β, α (learning rate), κ (emotion sensitivity in Equation 3.9) and Q_0 (to initialize $Q(s, a) = Q_0$ and $\rho(s) = Q_0$ for all s and a) in the affective RL algorithm

---

[1] these parameters were not tuned for optimizing performance.

were optimized to find the best-performing values. In addition, η was optimized in the Affective RL model with both experienced-based and prediction-based modes.

The simulation results[2] in Table 3.1 show the average and standard deviation of total rewards received during each phase.Each phase is composed of 1,000 steps in Chain and Loop, and of 20,000 steps in Maze. Also, the statistics for the affective RL algorithm were taken over 100 runs for Chain and Loop, and 10 runs for Maze.[3]. The results show that the affective RL algorithm beats other learning algorithms in terms of accumulated rewards.

Affective RL-based implementations performed better than other algorithms. Also, note that adding the prediction-based mode (affective RL with both modes) improved the performance significantly in the Maze problem, compared to the affective RL with only the experience-based mode. This is because the local confidence-based controls greatly helped the balance between exploration and exploitation in this large-scale domain (Maze).

---

[2]For comparison with other Q-learning based algorithms (Uniform, Boltzmann, Interval, IEQL+), we present the simulation results by (Dearden et al. 1998) tested with best-performing parameters for each algorithm.

[3]The statistics for other algorithms were taken over 10 runs for all three problems.

Chapter 4

Human Decision Experiments: Method and Hypotheses

Prospect theory (PT) was originally constructed to describe people's behavior in one-shot decision situations under risk (i.e., situations in which people make a one-shot choice with full knowledge of the outcome probability distributions of alternatives). Although some behavioral decision-making models and reinforcement learning models in computational learning theory have been applied to analyze people's behavior in decision-making situations under uncertainty, they have not fully incorporated the characteristics of PT-based subjective value functions (i.e., risk attitudes depending on reference-point dependency, diminishing sensitivity, loss-aversion) for modeling people's subjective experience and affective prediction. The proposed AC model uses PT-based parameterized subjective value functions to model people's experienced-utility and predicted-utility functions. The AC model hypothesizes that the shapes (parameters) of these subjective value functions dynamically vary with the decision-maker's task-related and/or incidental affective states in learning and decision making under uncertainty. Human decision-making experiments were conducted to empirically infer how people adjust the parameters (i.e., risk attitude and reference point) of their experienced-utility and predicted-utility functions in sequential decision-making situations involving incidental affective states (e.g., anger, fear) and task-related affects (e.g., confidence). This chapter presents the method and hypotheses of the experiments.

4.1 Experiment Method

Experiments were devised to observe human learning and decision making under uncertainty, frames and emotions. Basically there were four *emotion* conditions (neutral, anger, fear, economic fear) compared in a between-subject design. In each emotion condition, there were two *framing* conditions (gain frame and loss frame, between-subject) set by the experimenter. All respondents conducted two kinds of gambling tasks (Domain 1 and Domain 2, within-subject) involving different uncertain outcome distributions.

4.1.1 Respondents

Eighty-four respondents (49 males, 35 females) in the age range of 18-65 years (with a mean of 34 and a standard deviation of 12) participated in this experiment. Most respondents were recruited by advertisements offering a $50 gift card plus the chance to add $20 according to their achievement of goals in the experiment tasks (i.e., more incentive of $20 for the top 10% of the best decision makers who have won the most money in total over the tasks) in exchange for one hour of participation. The majority (about 70% of all respondents) were bachelor's degree holders (38%) or more advanced degree holders (32%).

4.1.2 Procedure

In the separate experiment-preparation room, the experimenter explained the purpose of the study (i.e., gathering data to develop a computational model of human decision-making) and asked each respondent to fill out a consent form. The respondent was also asked to wear a wristband-type skin conductance sensor on the non-dominant hand. This skin conductance sensor [45] developed at the Media Lab is a wristband with two Ag/AgCl electrodes to measures electrodermal activity (EDA) from the wrist at the sampling frequency of 32Hz and record the signal into an internal 2GB microSD card. No conductive gel was applied to the electrodes.

Then, each respondent was seated in a private cubicle (equipped with computers and headsets) in the experiment room and asked to follow the instructions on the computer. The computer in front of the respondent included a tiny built-in webcam taking video of

his or her face. When respondents launched the computer-based experiment program, they were assigned to one of the emotion conditions, based on the order that they signed up to participate in the study for practical purposes, which was random. After asking respondents to answer the profile survey (sex, age, education level, etc.), the program introduced a tutorial session for helping them understand the experiment tasks, and then started an actual experiment session. There were instructions before every step of the experiment session; thus, respondents were able to conduct all the experiment steps in the experiment session without any help or interruption of the experimenter. The experiment session in the neutral, anger and fear conditions is composed of the following steps:

1. Watching a neutral video clip (for inducing neutral affect as the baseline affect)

2. Answering the baseline measures of affect and the baseline financial survey

3. Writing a diary on a personal experience (congruent to the emotion condition)

4. Watching the first video clip (congruent to the emotion condition)

5. Performing the first decision task

6. Watching the second video clip (congruent to the emotion condition)

7. Performing the second decision task

8. Watching the third video clip (congruent to the emotion condition)

9. Answering the financial survey task on the banking attitude

10. Answering the post-experiment questionnaire (for emotion-manipulation check)

In the neutral, anger and fear conditions, respondents were asked to perform three main tasks in the following sequence, the first decision task - the second decision task - the financial survey task on the banking attitude, watching a video clip between any two of these main tasks. Yet, in the economic fear condition, respondents did these tasks in a different sequence, the financial survey task on the banking attitude - the first decision task - the second decision, still watching a video clip between any two of these main tasks. [1] In the analysis, we assumed that the sequence effect could be ignored due to the fact that respondents watched a video clip every time before starting a new task. Using the after-the-experiment questionnaire for manipulation checks, we checked that watching a video

[1]This was because we also wanted to compare the economic fear condition with other economic Ad conditions for another experimental purpose. Thus, the economic fear condition had to use a sequence that other economic Ad conditions employed.

clip reset respondents' emotional state to the target emotion for the emotion condition.

In the economic fear condition, the experiment session had the following steps:

1. Watching a neutral video clip (for inducing neutral affect as the baseline affect)

2. Answering the baseline affect measures and the baseline financial survey

3. Writing a diary on a personal experience (congruent to the emotion condition)

4. Watching the first video clip (congruent to the emotion condition)

5. Answering the financial survey on the banking attitude

6. Watching the second video clip (congruent to the emotion condition)

7. Performing the first decision task

8. Watching the third video clip (congruent to the emotion condition)

9. Performing the second decision task

10. Answering the post-experiment questionnaire (for emotion-manipulation check)

On every trial in a decision task (with total 25 trials), respondents were asked to select one of two options and to answer questions on the computer. Each decision task lasts approximately 10 minutes.

4.1.3 Instructions during the tutorial session

The following shows an example of the instructions during the tutorial session provided to a respondent given decision tasks in the gain frame:

This is a tutorial section for the gambling games you will complete during this study. You will have two INDEPENDENT gambling games in the actual study. For each game you will begin with an initial balance of $0. Each game includes 25 trials.

For each trial you will pay $20 and then select one of two options. Each selection will provide a random outcome (so you may win some money back).

It is very important to note that, in each game, one option has a higher average outcome than the other option. Your goal of each game is to figure out which option is better (in the long run) with fewer trials in order to minimize your total financial loss.

Also, please carefully answer the questions following each trial. Your opinions about your experience are really important in our data analysis.

For only the respondents who carefully answered their opinions, we will choose the top

10% of the best decision makers (who have won the most money in total over the games) and give an additional reward (up to $20) to them later. (We will email you later if you are in the top 10%.)

Note that when respondents were given the decision tasks in the loss frame, the following instruction was given instead of the one above: *For each trial you will receive $20 and then select one of two options. Each selection will provide a random outcome (so you may lose some money).*

4.1.4 Baseline affect measures

After the tutorial session, respondents started the experiment session. First, they watched a short (one-min) video clip "sticks" previously shown to induce neutral affect [49]. Then, they completed the form on the baseline measures of affect. Eight affective states ("angry", "anxious", "disgusted", "fearful", "happy", "interested", "irritated", "sad") were included on the form. Response scales ranged from 0 (do not feel the emotion the slightest bit) to 8 (feel the emotion even more strongly than ever before) for each affective state [35].

4.1.5 Baseline financial survey

Respondents were asked to answer the baseline financial survey on the banking services and the prospects of US economy as follows:

(Q1) Please think about the banking and financial services industry as a whole. How would you rate your overall impression of the banking industry?

Poor								Excellent
1	2	3	4	5	6	7	8	9

(Q2) Using a scale of 1 to 5, where 1 is "Poor" and 5 is "Excellent", how would you rate the health of the U.S. economy today?

Poor				Excellent
1	2	3	4	5

(Q3) Now, thinking about how today's economy compares to the past, using the scale below how would you compare today's economy to as it was...

Compared to this time last year, today's economy is:

much worse	a little worse	the same	a little better	much better

Compared to this time 3 years ago, today's economy is:

much worse	a little worse	the same	a little better	much better

(Q4) Thinking about where you expect the economy to be in the future relative to today, using the scale below, how would you expect the economy to be...

Compared to today's economy, this time next year would be:

much worse	a little worse	the same	a little better	much better

Compared to today's economy, this time in 3 years would be:

much worse	a little worse	the same	a little better	much better

4.1.6   Emotion manipulations

To elicit the emotional state associated with each emotion condition, respondents were asked to write about a personal experience in five minutes, and then they were guided to watch a short video clip depending on the emotion condition (neutral, anger, fear, economic fear).

106

Writing a diary on a personal experience

As for the diary part, respondents in the neutral condition were asked to write about their daily activities. Respondents in the anger (or fear) condition were asked to write about one situation that made them angry (or fearful). Also, respondents in the economic fear condition were asked to write about one event that makes them most fearful about the economy (in the economic fear condition). The following introduction messages were presented for each emotion condition:

(Neutral condition) Please describe, as best you can, how you typically spend your evenings. You might begin by writing down a detailed description of your activities, and then figure out how much time you devote to each activity. If you can, please write your description so that someone reading this might be able to reconstruct your typical routine.

(Anger condition) Please describe, as best you can, one situation that made you ANGRY. In particular, please write about a situation in which you experienced significant injustice or humiliation, such as arrogant/offensive behavior of your partner/neighbors or bosses, racial/sexual discrimination, verbal/physical abuse, oppressive social system or irrational government policy. This could be a situation you are presently experiencing or something from the past. Begin by writing down what you remember of the ANGER-INDUCING EVENT, and continue by writing as detailed a description of the event as is possible. If you can, write your description so that someone reading it might get angry from learning about the situation. What is it like to be in this situation? Why does it make you so ANGRY?

(Fear condition) Please describe, as best you can, one situation that made you FEARFUL. In particular, please write about a situation in which you experienced significant threat, such as the threat of danger, death, or significant loss or the threat of social rejection or humiliation. According to surveys, some of the most commonly feared situations are enclosed spaces, tunnels and bridges, social rejection, home intruders, failure, terror, and public speaking. This could be a situation you are presently experiencing or something from the past. Begin by writing down what you remember of the FEAR-INDUCING EVENT, and continue by writing as detailed a description of the event as is possible. If you can, write your description so that someone reading it might get fearful from learning about the situation. What is it like to be in this situation? Why does it make you so FEARFUL?

(Economic Fear condition) The recent economic recession evoked a lot of emotion in Americans. Many people have lost their livelihoods, jobs, homes, and the economists do not agree on whether

or when things will get better. We are particularly interested in what makes you most FEARFUL about the economy. Please describe in detail the one thing that makes you most fearful about these events. Write as detailed a description of that thing as possible. If you can, write your description so that someone reading it might get fearful from learning about the situation. What aspect of the economic recession makes you the most fearful? What is the worst thing you can imagine that would happen to you financially? (e.g. loss of ability to work, being sued for all you own ...) Why does it make you so FEARFUL?

Watching the "first" video clip for eliciting the target emotion

For the video part, respondents were asked to put on the headset and watch video clips according to the emotion condition. Each video clip lasted approximately 3~4 minutes. Before watching a video clip, respondents were given some information on the video clip as well as the following messages: *As part of our interest in how people respond to visual information, we want you to be absorbed in the video. We will ask you some questions about the video and your experience watching it later in the study.* In addition, respondents in the anger, fear and economic fear conditions were asked to *imagine how they would feel if they were experiencing the situation portrayed.*

Respondents in the neutral condition watched a video clip about a natural park in Alaska showing peaceful and calm scenes by Alaska Channel.

Respondents in the anger condition watched a video clip from the movie "The Bodyguard (1980)" showing a bully who humiliated and beat up a teenager. Before watching the video clip, they were given the following information: "The bully was charged for his crime and found not guilty because of a technicality. The bully and his friend walked away from the trial as free men, and both have been in trouble with the law subsequently."

Respondents in the fear condition watched a scary video clip from the movie "The Silence of the Lambs (1991)" showing a woman pursuing a psychopath killer who skinned his female victims. Prior research has shown that these video clips employed for anger and fear conditions are an effective means of eliciting the target emotions [49].

Respondents in the economic fear condition watched a video clip in which the US economic crisis (foreclosure, credit crunch, unemployment, etc.) was shown with an economic

expert's talk on the gloomy prospects of the US economy. Before watching the video clip, they were asked to imagine how they would feel if they were experiencing one or more miserable economic situations portrayed in the video clip.

After watching the "first" video clip, respondents were asked to perform their first decision task (the neutral, anger and fear conditions) or answer the financial survey on the banking attitude (the economic fear condition).

Watching the "second" video clip for eliciting the target emotion

Before performing the second decision task (the neutral, anger and fear conditions) or performing the first decision task (the economic fear condition), respondents watched another video clip for eliciting the target emotion of their emotion condition. Considering possible influences of experiences during the first decision task on respondents' emotional state, we added this step for resetting respondents' emotional state to the target emotion.

Respondents in the neutral condition watched a video clip showing fish at the Great Barrier Reef from a National Geographic Special [35].

Respondents in the anger condition watched a video clip from the movie "Cry Freedom (1987)" showing the police raiding a black township along with killing and wounding many residents. Before watching the video clip, they were given the following information: "This movie clip is from the film based on the true story of Steve Biko, the Black Consciousness Movement leader during the apartheid era of South Africa, who was arrested and killed while in police custody."

Respondents in the fear condition watched a scary video clip from the movie "The Shining (1980)" showing a little boy walking in a hallway. Prior researches have shown that these video clips employed for anger and fear conditions are an effective means of eliciting the target emotions [49].

Respondents in the economic fear condition watched an ABC news clip on the story of scared Americans in tough times of the US economy titled "Tough times".

After watching the "second" video clip, respondents were asked to perform their second decision task.

Watching the "third" video clip for eliciting the target emotion

Before answering the financial survey on the banking attitude (the neutral, anger and fear conditions), respondents in neutral, anger and fear conditions watched the "first" video clip used for eliciting the target emotion once again, which they had watched immediately before starting their first decision task.

Before performing the second decision task (the economic fear condition), respondents in the economic fear condition watched an ABC news clip on the falling US economy and unemployment situations titled "Falling Fast". Considering possible influences of experiences during the second decision task on respondents' emotional state, we added this step for resetting respondents' emotional state to the target emotion.

After watching the "third" video clip, respondents were asked to answer the financial survey and the post-experiment questionnaire.

4.1.7 Decision tasks

The decision tasks under uncertainty were devised to infer how the respondent's experience and prediction influence their decision behavior. Each respondent played the tasks of two domains (Domain 1 and Domain 2, explained below) in random order. Each task was a two-armed bandit task designed for evaluating the respondent's risk attitude: there are two options whose outcome distributions are unknown to the respondent. The respondent has 25 trials during each task and the goal is to maximize the total outcome (or minimize the total financial loss) over all trials. For each trial the respondent will receive \$20 (or pay \$20) and then select one of two options. Each selection will provide a random outcome so the respondent may lose (or win) some money.

Domain 1 and Domain 2: One option is risky (involving high variance in outcomes) and the other option is safe (involving low variance in outcomes). There were two domains (underlying outcome distributions) used in our experiment. For Domain 1, the safe option is the optimal option (with a greater outcome on average) and the risky option is the suboptimal option (with a smaller outcome on average). For Domain 2, the risky option is the optimal choice and the safe option is the sub-optimal choice.

110

Gain and Loss frames: For almost the half (11 respondents), the decision tasks were given in terms of the gain frame, i.e., on each trial they first paid some amount of bet money and then won a random amount of money according to their choice. For the other half (10 respondents), the decision tasks were given in terms of the loss frame, i.e., on each trial they received some amount of bet money and then lost a random amount of money according to their choice.

- Gain frame:

  Please click the Bet button to start a new trial (this is trial XX of 25). (By clicking, you pay $20 for this bet, but you might win it back.)

- Loss frame:

  Please click the Bet button to start a new trial (this is trial XX of 25). (By clicking, you receive $20 for this bet, but you might lose it.)

Actual outcome distributions used in the experiment

In each domain, the safe option had a Gaussian outcome distribution with a small standard deviation (= $2) whereas the risky option had its outcome distribution as a mixture of three Gaussian distributions with different modes.

Gain frame on Domain 1:

- Safe-option outcome distribution $\sim N(\text{mean} = \$22, \text{std} = \$2)$

- Risky-option outcome distribution $\sim 0.4\ N(\text{mean} = \$30, \text{std} = \$2) + 0.2\ N(\text{mean} = \$18, \text{std} = \$2) + 0.4\ N(\text{mean} = \$6, \text{std} = \$2)$

The risky-option outcomes were sampled from a very high mode $N(\text{mean} = \$30, \text{std} = \$2)$ and a very low mode $N(\text{mean} = \$6, \text{std} = \$2)$ with 40% probability each and from a mid mode $N(\text{mean} = \$18, \text{std} = \$2)$ with 20% probability. Note that in this case (Domain 1, gain frame), the average outcome of the safe option over trials (approximately $22) is greater than that of the risky option (approximately $18).

111
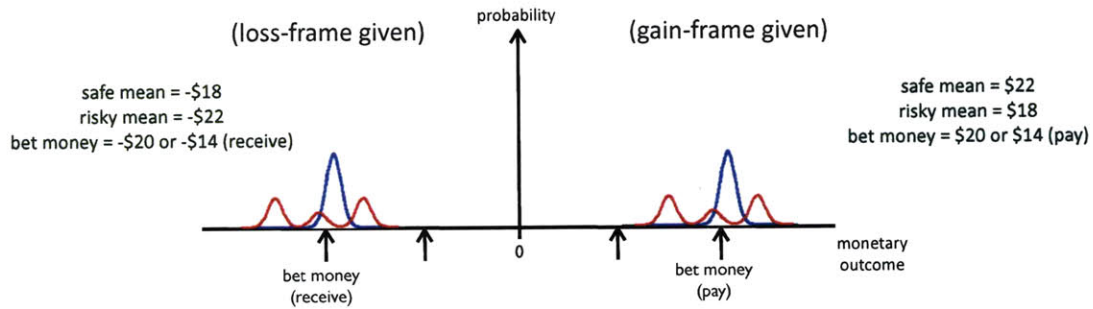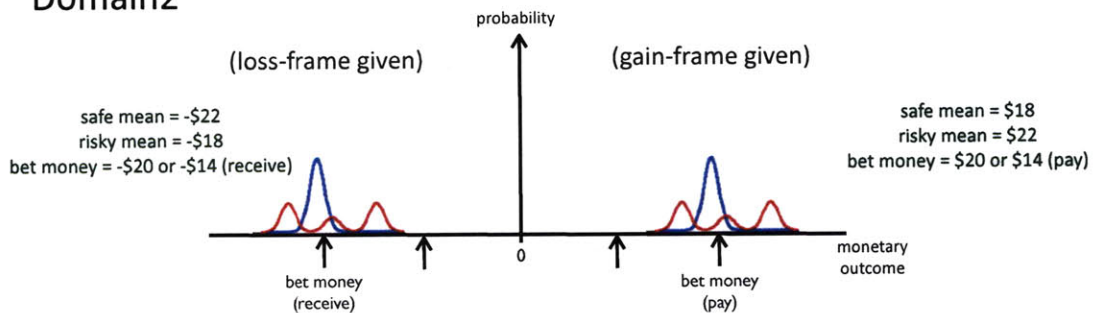
Domain 1

Domain2

Figure 4-1: Four different cases for outcome probability distributions of two options (Gain frame on Domain 1, Loss frame on Domain 1, Gain frame on Domain 2, Loss frame on Domain 2). Each respondent played the tasks of two domains (Domain 1 and Domain 2) in random order given in terms of either the gain frame or the loss frame.

Loss frame on Domain 1:

- Safe-option outcome distribution $\sim N(\text{mean} = -\$18, \text{std} = \$2)$

- Risky-option outcome distribution $\sim 0.4\ N(\text{mean} = -\$10, \text{std} = \$2) + 0.2\ N(\text{mean} = -\$22, \text{std} = \$2) + 0.4\ N(\text{mean} = -\$34, \text{std} = \$2)$

That is, the risky-option outcomes were sampled from a very high mode $N(\text{mean} = -\$10, \text{std} = \$2)$ and a very low mode $N(\text{mean} = -\$34, \text{std} = \$2)$ with 40% probability each and from a mid mode $N(\text{mean} = -\$22, \text{std} = \$2)$ with 20% probability. Note that in this case (Domain 1, loss frame), the average outcome of the safe option over trials (approximately -$18) is greater than that of the risky option (approximately -$22).

112

Gain frame on Domain 2:

- Safe-option outcome distribution $\sim N$(mean = \$18, std = \$2)

- Risky-option outcome distribution $\sim 0.4\ N$(mean = \$34, std = \$2) + 0.2 N(mean = \$22, std = \$2) + 0.4 N(mean = \$10, std = \$2)

In other words, the risky-option outcomes were sampled from a very high mode N(mean = \$34, std = \$2) and a very low mode N(mean = \$10, std = \$2) with 40% probability each and from a mid mode N(mean = \$22, std = \$2) with 20% probability. Note that in this case (Domain 2, gain frame), the average outcome of the risky option over trials (approximately \$22) is greater than that of the safe option (approximately \$18).

Loss frame on Domain 2:

- Safe-option outcome distribution $\sim N$(mean = -\$22, std = \$2)

- Risky-option outcome distribution $\sim 0.4\ N$(mean = -\$6, std = \$2) + 0.2 N(mean = -\$18, std = \$2) + 0.4 N(mean = -\$30, std = \$2)

That is, the risky-option outcomes were sampled from a very high mode N(mean = -\$6, std = \$2) and a very low mode N(mean = -\$30, std = \$2) with 40% probability each and from a mid mode N(mean = -\$18, std = \$2) with 20% probability. Note that in this case (Domain 2, loss frame), the average outcome of the risky option over trials (approximately -\$18) is greater than that of the safe option (approximately -\$22).

In our experiment, we actually employed fixed sequences of random outcomes from each outcome distribution rather than sampling outcomes for each respondent. That is, the sequence of experienced outcomes for each option is the same across respondents for a given domain and frame condition. Appendix D lists the fixed sequences of random outcomes used in the experiments. In addition, for the risky option, we had two kinds of fixed outcome sequences: one sequence starting with a very high outcome (from a very high mode) and the other sequence starting with a very low outcome (from a very low mode) on the first trial of the risky option. Almost half of the respondents in each emotion condition were given a very high outcome on their first trial of the risky option, and the rest in the condition were given a very low outcome.

113

Self-reported experienced utility (EU), confidence, and predicted utility (PU)

After obtaining an outcome, respondents answered three kinds of self-reported measures in sequence on each trial: experienced utility (self-report 1), confidence (self-report 2), and predicted utilities (self-report 3).

Self-report 1 (Experienced utility (EU)): Respondents self-reported their experienced pleasure or displeasure (called the "experienced utility") for the obtained outcome, answering the following question (See Figure 4-2):

(Gain frame) *In this trial you PAID $XX. But you WON $XX by choosing Option YY. How pleased or displeased are you with the obtained outcome? (click on the bar below and adjust the handle that will appear.)*

(Loss frame) *In this trial you RECEIVED $XX. But you LOST $XX by choosing Option YY. How pleased or displeased are you with the obtained outcome? (click on the bar below and adjust the handle that will appear.)*

Figure 4-2: Asking the experienced utility (an example of the gain frame)

Figure 4-3: Asking the confidence in the next choice

Figure 4-4: Asking the predicted utility of each option

Self-report 2 (Confidence): Respondents self-reported their confidence in the next choice in terms of achieving the goal of the task, answering the following question (See Figure 4-3):

How confident are you that your next choice will result in minimizing your total financial loss? (click on the bar below and adjust the handle that will appear.)

Self-report 3 (Predicted Utility (PU)): Respondents self-reported their prediction on the future experienced pleasure or displeasure (called "predicted utilities") with the outcome obtained after selecting each option (Option 1 or Option 2) on the next trial, answering the following question (See Figure 4-4):

Based on your overall experience with each option (if you had any), how pleased or displeased will you be with the outcome if you select Option 1 (or Option 2) in your next trial?

4.1.8 Financial survey on the banking attitude

Respondents answered the following financial survey on the banking attitude:

(Q1) If you were going to open a new account or obtain a new loan, how likely would you be to open a new account or obtain a new loan with Bank of America? (1 10 scale: 1= not at all likely, 10 = extremely likely)

(Q2) If you were going to open a new investment account such as a brokerage account, mutual fund or IRA, how likely would you be to open a new investment account with Bank of America? (1 10 scale: 1= not at all likely, 10 = extremely likely)

Please rate your agreement with the following statements.

(Q3): Overall, Bank of America does not make misleading claims. (1 10 scale: 1 = poor, 10 = excellent)

(Q4): Overall, Bank of America keeps promises it makes to customers. (1 10 scale: 1 = poor, 10 = excellent)

(Q5): Overall, Bank of America is trustworthy. (1 10 scale: 1 = poor, 10 = excellent)

(Q6): Overall, Bank of America is financially stable. (1 10 scale: 1 = poor, 10 = excellent)

(Q7): Overall, Bank of America is a safe place to keep my money. (1 10 scale: 1 = poor, 10 = excellent)

4.2 Experiment Hypotheses

In each trial, after obtaining an outcome, respondents answered three kinds of self-reported measures in sequence: experienced utility (self-report 1), confidence (self-report 2), and predicted utilities (self-report 3). We infer two distinct subjective utility functions: The experienced-utility (EU) function for evaluating the obtained outcomes is inferred from the self-report 1. The predicted-utility (PU) function for predicting the future experience is inferred from the self-reports 2 and 3.

Figure 4-5 shows the experiment design of human decision making under uncertainty (domains), frames and emotions. In this design, the (used) EU and PU frames can be viewed as a hidden latent variable inferred from respondents' self-reports.
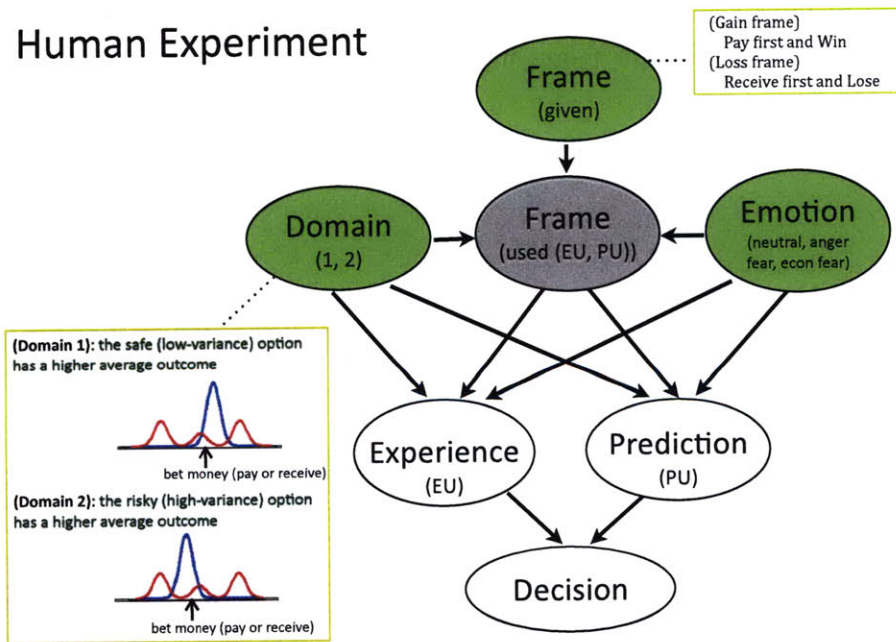
Figure 4-5: Asking the experienced utility (an example of the gain frame)

The data analysis will be based on three different kinds of frames: the given frame, the

EU frame (i.e., the frame the decision maker actually employed to evaluate experienced-utility), and the PU frame (i.e., the frame the decision maker actually employed to estimate predicted-utility). The given frame is set by the experimenter. The EU frame is associated with the reference point of the EU function computed by the data-fitting analysis of respondents' self-reported experiences during the decision task. The PU frame is associated with the reference point of the PU function computed by the data-fitting analysis of respondents' self-reported confidences and predictions during the decision task.)

4.2.1 The analysis based on the give frame

H1. The frame (gain or loss frame given) and the domain (1 or 2) influences behavioral choices under uncertainty.

[H1-1] Respondents select the risky option more often on Domain 2, where the risky option has a higher average of outcomes than the safe option, than on Domain 1, where the safe option has a higher average of outcomes than the risky option.

[H1-2] On both Domain 1 and Domain 2, respondents employing the loss frame select the risky option more often than those employing the gain frame. That is,

On Domain 1 (where the safe option has a higher average of outcomes than the risky option), respondents select the safe option more often in the gain frame than in the loss frame.

On Domain 2 (where the risky option has a higher average of outcomes than the safe option), respondents select the risky option more often in the loss frame than in the gain frame.

H2. The frame (gain or loss frame given) and the domain (1 or 2) influence subjective discriminability. First, objective discriminability is defined as

$$d'_{obj} = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

where

μ_1 = the average of obtained outcomes from the optimal option (i.e., the safe option on Domain 1 or the risky option on Domain 2) over trials,

118

μ_2 = the average of obtained outcomes from suboptimal option (i.e., the risky option on Domain 1 or the safe option on Domain 2) over trials,

σ_1 = the standard deviation of obtained outcomes from the optimal option over trials,

σ_2 = the standard deviation of obtained outcomes from the suboptimal option over trials.

The definition of objective discriminability depends on a decision maker's obtained outcomes over trials. That is, the measure of objective discriminability relies only on the underlying outcome distributions and samplings from them, but should not be affected by the gain or loss frame used to evaluate experienced utilities.

Second, subjective discriminability is defined as

$$d'_{subj} = \frac{\mu_{1,eu} - \mu_{2,eu}}{\sqrt{\sigma^2_{1,eu} + \sigma^2_{2,eu}}}$$

where

$\mu_{1,eu}$ = the average of experienced utilities from the optimal option over trials,

$\mu_{2,eu}$ = the average of experienced utilities from suboptimal option over trials,

$\sigma_{1,eu}$ = the standard deviation of experienced utilities from the optimal option over trials,

$\sigma_{2,eu}$ = the standard deviation of experienced utilities from the suboptimal option over trials.

In addition, we define the discriminability ratio as the ratio of subjective discriminability over objective discriminability:

d' ratio (discriminability ratio) = d'_{subj}/d'_{obj}

The definition of subjective discriminability depends on the decision maker's experienced utilities measured as self-reported values of the experienced pleasure/displeasure after obtaining outcomes) over trials. Thus, subjective discriminability is an effective measure to see how a respondent's subjective experiences of previously obtained choice outcomes influence their future choice selection under uncertainty.

A larger subjective discriminability makes respondents more likely to choose the optimal option, whereas a smaller subjective discriminability makes respondents more likely to choose the sub-optimal option. Also, a positive value of subjective discriminability indi-

119

cates that experiences generally contribute to selecting the optimal option more, whereas a negative value means the opposite tendency, selecting the suboptimal option.

[H2-1] On Domain 1, subjective discriminability in the gain frame is significantly greater than that in the loss frame while objective discriminabilities of the two frames do not differ significantly.

[H2-2] On Domain 2, subjective discriminability in the loss frame is significantly greater than that in the gain frame while objective discriminabilities of the two frames do not differ significantly.

[H2-3] On Domain 1, respondents in the gain frame have a greater subjective discriminability than the objective discriminability, while respondents in the loss frame have a smaller subjective discriminability than the objective discriminability.

[H2-4] On Domain 2, respondents in the loss frame have a greater subjective discriminability than the objective discriminability, while respondents in the gain frame have a smaller subjective discriminability than the objective discriminability.

4.2.2 The analysis based on the EU frame

H3. Emotion conditions influence the reference point and the shape of the experienced-utility function.

Inferring the experienced-utility function: For each emotion condition, the experienced-utility function over respondents was inferred from the experienced-utility responses (self-report 1). The mixed-effect model fitting for EU (experienced-utility) assumed that the shape parameters were considered fixed effects and the reference-point parameters (of two domains) were random effects. The mixed-effect EU model returns the reference points (for experience) each respondent employed during two tasks can be inferred ($x_{ref1,j}$ (domain 1) and $x_{ref2,j}$(domain 2) for respondent j). The reference point for experience (or the EU frame) is determined by this estimation of the experienced-utility function and its reference points.

The EU frame is defined as follows. We first compute the average outcome of the safe (low-variance) option. Then we test if the reference point the respondent used for experienced utility (self-report 1) is larger or smaller than that average outcome. If it is

larger, then the EU frame is a loss frame; else it is a gain frame. Note that we could compute this after each trial, but here it is computed only once, aggregating all trials.

H4. The domain, the EU frame and the emotion condition influence the discriminability ratio $(= d'_{subj}/d'_{obj})$.

[H4-1] On Domain 1 (where optimal option = safe option), respondents across all emotion conditions will show greater discriminability ratios (leading to more risk-aversion and optimal selections) in the EU gain frame than in the EU loss frame.

[H4-2] On Domain 2 (where optimal option = risky option), respondents across all emotion conditions will show greater discriminability ratios (leading to more risk-seeking and optimal selections) in the EU loss frame than in the EU gain frame.

[H4-3] On each domain, the emotion condition influences the discriminability ratio, depending on the EU frame (gain or loss frame).

H5. Experience, gender and emotion influence confidence, which influences future prediction and decision.

[H5-1] When respondents have a relatively good experience, they get more confident in the task, compared to when they had a relatively bad experience.

[H5-2] When respondents are more confident in the task, they tend to predict that one option is much better than the other, compared to when they are less confident.

[H5-3] When respondents tend to predict that one option is better than the other, they are more likely to select the better option (= exploitative choice).

There are the gender difference on risk estimates. Lerner et al. [33] observed that males had less pessimistic risk estimates than did females, emotion differences explaining 60 to 80% of the gender difference. Here it is hypothesized that there are interaction effects of the gender and the emotion condition on respondents' self-reported confidence and probability of selecting the exploitative choices.

[H5-4] The gender and the emotion condition influence confidence and exploitative-exploratory choice behavior.

4.2.3 The analysis based on the PU frame

H6. Emotion conditions influence the reference point and the shape of the predicted-utility function.

Inferring the predicted-utility function: For each emotion condition, the predicted-utility function over respondents was inferred from the predicted-utility and confidence responses (self-reports 2 and 3). We used a mixed-effect model to infer the predicted-utility function of each emotion condition, with the shape parameters considered fixed effects and the reference-point parameters (of two domains) considered random effects. The mixed-effect PU model returns the reference points (for prediction) that each respondent employed during two tasks ($x_{ref1,j}$ (domain 1) and $x_{ref2,j}$ (domain 2) for respondent j). The reference point for prediction (or the PU frame) is determined by this estimation of the predicted-utility function and its reference points.

The PU frame is defined as follows. We first compute the average outcome of the safe (low-variance) option. Then we test if the reference point the respondent used for predicted utility (self-report 3) is larger or smaller than that average outcome. If it is larger, then the PU frame is a loss frame; else it is a gain frame. Note that we could compute this after each trial, but here it is computed only once, aggregating all trials.

H7. The domain, the PU frame and the emotion condition influence the probability of choosing the risky option.

Note that this hypothesis is associated with the behavioral choice measure.

[H7-1] On both Domain 1 and Domain 2, respondents employing the PU loss frame select the risky option more often than those employing the PU gain frame.

[H7-2] On each domain, the emotion condition influences the probability of choosing the risky option, depending on the PU frame (gain or loss frame).

H8. The domain, the PU frame and the emotion condition influence the average PU difference of the risky option and the safe option (= Average[PU(risky option)] - Average[PU(safe option)]).

Note that this hypothesis is associated with the self-reported PU measure.

By this definition, when the average PU difference (PUrisk - PUsafe = Average[PU(risky option)] - Average[PU(safe option)]) is positive, the average PU of the risky option is greater than that of the safe option by the definition. Otherwise, the average PU of the safe option is greater than that of the risky option. Thus, the average PU difference (PUrisk - PUsafe) can serve as a measure of risk attitude in terms of the predicted utilities.

[H8-1] On both Domain 1 and Domain 2, respondents employing the PU loss frame have a greater value of the average PU difference (leading to more risk seeking) than those employing the PU gain frame.

[H8-2] On each domain, the emotion condition influences the probability of choosing the risky option, depending on the PU frame (gain or loss frame).

Chapter 5

Human Decision Experiments: Analysis and Results

This chapter explains the analysis and results of human decision experiments to confirm main hypotheses described in Chapter 4. The AC model will be applied to infer the experienced-utility and predicted-utility functions from respondents' self-reports in each emotion condition. The data analysis is based on three different kinds of frames: the given frame, the EU frame (i.e., the frame the decision maker actually employed to evaluate experienced utility), and the PU frame (the frame the decision maker actually employed to estimate predicted utility). The given frame-based analysis will show the main effect of the given frame (gain or loss frame set by the experimenter) on behavioral choices (risk averse or risk seeking) and subjective discriminability. In addition, the EU and PU frame-based analysis will clearly show the interaction effects of framing and emotion on human experience, prediction and decisions under uncertainty.

5.1 The analysis based on the give frame

The analysis based on the given frame will show how the frame condition and the domain type influence behavioral choices under uncertainty and subjective discriminability.

(a) domains   (b) frames
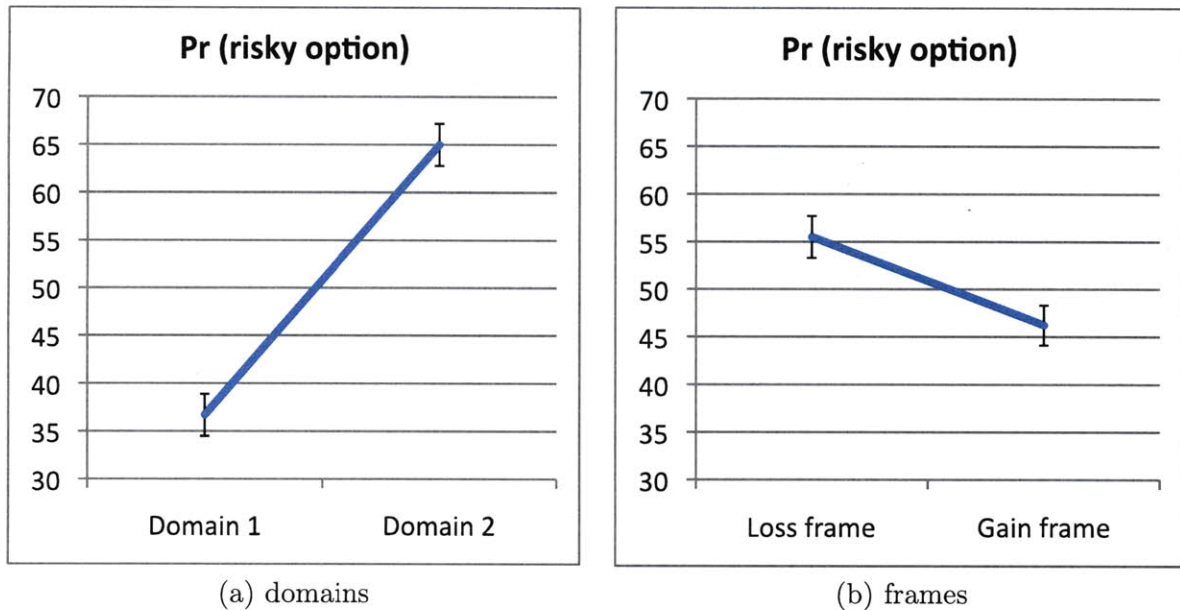
Figure 5-1: The main effect of domain types (a) or frame types (b) on the probability of selecting the risky option over trials (pooling over respondents in all four emotion conditions). Note that error bars indicate standard errors of the mean.

5.1.1 The frame (gain or loss frame given) and the domain (1 or 2) influences behavioral choices under uncertainty

A general linear model (GLM) three-factor ANOVA test was conducted to assess the influence of the domain type (domain 1, domain 2), the frame type (gain or loss frame given), and the emotion condition (neutral, anger, fear, economic fear) on the probability of selecting the risky option over all trials during a task.

There were significant main effects of the domain type ($M_{domain1} = 37\%$ vs. $M_{domain2} = 65\%$; $F(1,167) = 84.6$, $p < .001$) (supporting H1-1) and the frame type ($M_{gain} = 46\%$ vs. $M_{loss} = 56\%$; $F(1,167) = 9.53$, $p < .01$) on the probability of choosing the risky option, but no significant interaction between the domain type and the given frame type.

On Domain 1, respondents selected the optimal option (= the safe option) marginally more often in the gain domain than in the loss domain ($M_{gain} = 68\%$ vs. $M_{loss} = 59\%$; $F(1, 83) = 3.89$, $p = .052$). Also, on domain 2, respondents selected the optimal option (= the risky option) significantly more often in the loss frame than in the gain frame ($M_{gain} = 60\%$ vs. $M_{loss} = 70\%$; $F(1, 83) = 5.48$, $p < .05$) (supporting H1-2).

126
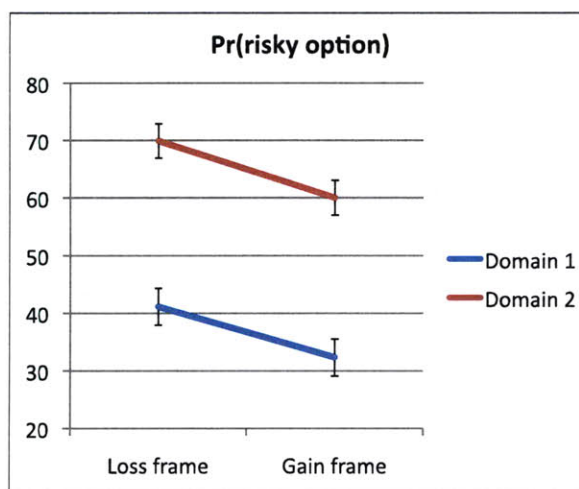
Figure 5-2: The interaction effect of the domain type and the frame type on the probability of selecting the risky option over trials (pooling over respondents in all four emotion conditions). Note that error bars indicate standard errors of the mean.

5.1.2 The frame (gain or loss frame given) and the domain (1 or 2) influences subjective discriminability

Pooling behavioral choice data over respondents across emotion conditions, on each domain, the frame type significantly influenced respondents' subjective discriminability when the objective discriminability was not significantly different between frame types. Excluding the tasks where respondents did not try each option for at least two trials, ANOVA tests showed the following:

On Domain 1, subjective discriminability in the gain frame was significantly greater than that in the loss frame ($M_{gain} = .52$ vs. $M_{loss} = .04$; $F(1,75) = 26.2$, $p < .001$) while objective discriminabilities of two frames did not differ significantly ($M_{gain} = .29$ vs. $M_{loss} = .32$; $F(1,75) = 1.7$, NS) (supporting H2-1).

On Domain 2, subjective discriminability in the loss frame was significantly greater than that in the gain frame ($M_{gain} = .20$ vs. $M_{loss} = .55$; $F(1,81) = 14.9$, $p < .001$) while objective discriminabilities of two frames did not differ significantly ($M_{gain} = .33$ vs. $M_{loss} = .32$; $F(1,81) = .25$, NS) (supporting H2-2).

On Domain 1, respondents in the gain frame had a significantly greater subjective discriminability than the objective discriminability ($M_{subj} = .52$ vs. $M_{obj} = .29$; $T(38) =$

127

Domain 1				Domain 2			
Gain frame		Loss frame		Gain frame		Loss frame	
d'_{subj}	d'_{obj}	d'_{subj}	d'_{obj}	d'_{subj}	d'_{obj}	d'_{subj}	d'_{obj}
.52	.29	.04	.32	.20	.33	.55	.32
$T(38)=3.66$		$T(36)= -3.90$		$T(42)= -2.57$		$T(38)= 3.42$	
$p < .001$		$p < .001$		$p < .01$		$p < .001$	

Table 5.1: Comparisons of the subjective discriminability with the objective discriminability in gain and loss frames on each domain type (pooling over respondents in all four emotion conditions).

(a) Domain 1   (b) Domain 2
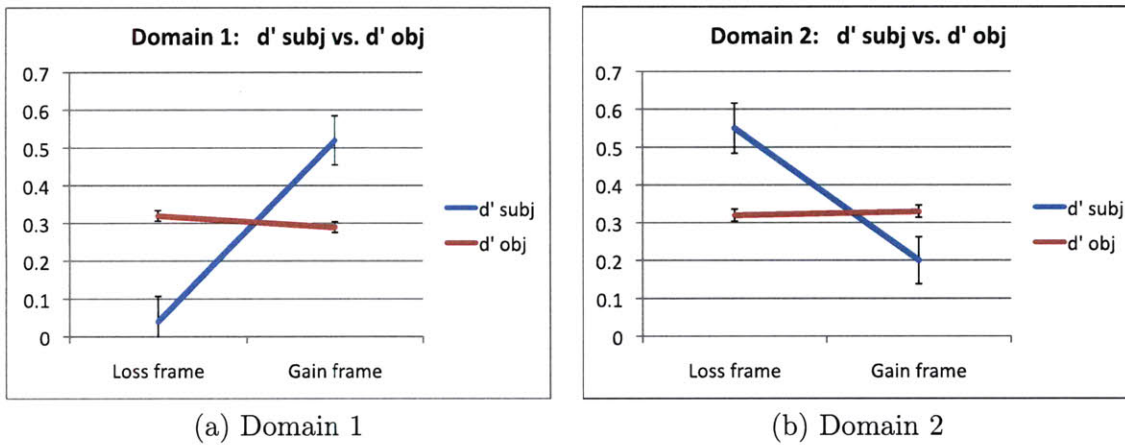
Figure 5-3: The subjective discriminability and the objective discriminability in gain and loss frames on each domain type (pooling over respondents in all four emotion conditions). Note that error bars indicate standard errors of the mean.

3.66, $p < .001$), while respondents in the loss frame have a significantly smaller subjective discriminability ($M_{subj} = .04$ vs. $M_{obj} = .32$; $T(36) = -3.90$, $p < .001$) (supporting H2-3).

On Domain 2, respondents in the loss frame have a greater "positive" subjective discriminability than the objective discriminability ($M_{subj}= .55$ vs. $M_{obj} = .32$; $T(38) = 3.42$, $p < .001$), while respondents in the gain frame have a significantly smaller subjective discriminability ($M_{subj} = .20$ vs. $M_{obj} = .33$; $T(42) = -2.57$, $p < .01$) (supporting H2-4).

Table 5.1 compares the subjective discriminability with the objective discriminability in each of the four pairs of the domain type (domain 1, domain 2) and the frame type (gain or loss frame given) across emotion conditions.

According to the definition of subjective discriminability, higher subjective discrim-

inability should be associated with higher probability of selecting the optimal option on each domain. This was confirmed by the behavioral choice data (Figure 5-2) and the subjective discriminability data (Figure 5-3). That is, the frame involving higher subjective discriminability on each domain (the gain frame on Domain 1 and the loss frame on Domain 2) led to more optimal choices: On Domain 1 the gain frame led to more safe (optimal) choices that resulted in higher average outcome. On Domain 2 the loss frame led to more risky (optimal) choices that resulted in higher average outcome.

5.2   The analysis based on the EU frame

The EU frame is the frame the decision maker actually employed to evaluate experienced utility. The EU frame can be inferred from the data-fitting analysis of respondents' self-reported experiences during the decision task. The EU frame-based analysis presents the interaction effects of framing and emotion on human experience and decisions. Figure 5-4 shows some examples of the EU gain frame and the EU loss frame. Note that even though (objective) outcomes are negative, they can be viewed as gains if the reference point is smaller than the outcomes. Similarly, even though (objective) outcomes are positive, they can be viewed as losses if the reference point is greater than the outcomes.

5.2.1   Emotion conditions influence the EU frame and the shape of the experienced-utility function

Respondents tended to use the bet money as their EU reference point (i.e., the reference point of their EU function estimated by the EU data fitting) in the (given) gain frame on Domain 1, in the loss frame on Domain 2, and in the gain frame on Domain 2. In other words, except in the loss frame on Domain 1, respondents in all emotion conditions tended to involve the EU gain frame when the bet money is smaller than the average outcome of the safe option, and involve the EU loss frame when the bet money is greater than the average outcome of the safe option.

Note that for the loss frame on Domain 1, respondents first received the bet money of $14 or $20 and then lost some money after selecting an option on each trial. Also, the

Gain EU frame

mean of the safe option > ref point

(tend to view outcomes of the safe option as gains)
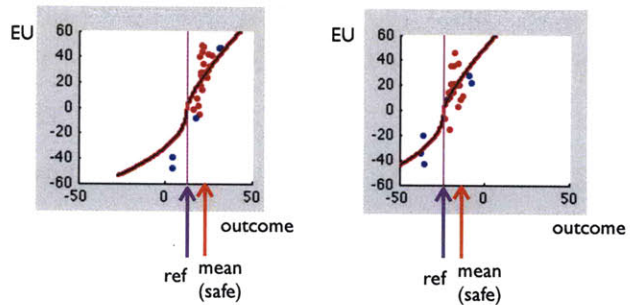
Loss EU frame

mean of the safe option < ref point

(tend to view outcomes of the safe option as losses)
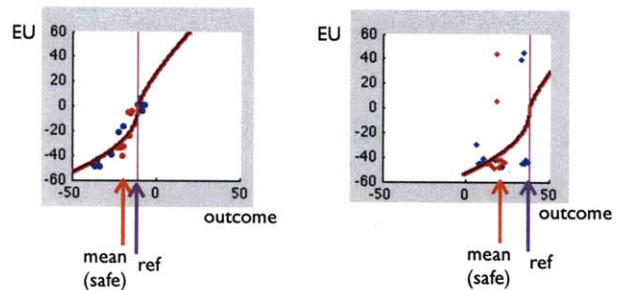
Figure 5-4: Examples of the EU gain frame and the EU loss frame: red dots - selecting the safe option, blue dots - selecting the risky option

average outcome of the safe option was a loss of $18.

For the loss frame on Domain 1, respondents in the neutral condition tended to use the bet money as their EU reference point. In other words, neutral respondents involved the EU loss frame when they received the bet money of $14 (i.e., self-reporting displeasure for the average loss of $18 from the safe option), but involved the EU gain frame when they received the bet money of $20 (i.e., self-reporting pleasure for the average loss of $18 from the safe option). That is, they tended to be very rational in light of the net outcome. Yet, *respondents in the negative-emotion (anger, fear, economic fear) conditions tended to employ the EU loss frame* self-reporting displeasure for the average loss of $18 from the safe option regardless of receiving $14 or $20 initially (Fisher's exact tests, $p = .02$ (anger), $p < .01$ (fear), $p = .06$ (economic fear)); their EU reference point tended to be greater than the average outcome of the safe option, so they view the obtained outcomes from the safe option as losses regardless of their bet-money condition.

Now the following analysis presents the inferred experienced-utility function from self-reported EUs in each emotion condition. Appendix B describes the EU data-fit method in
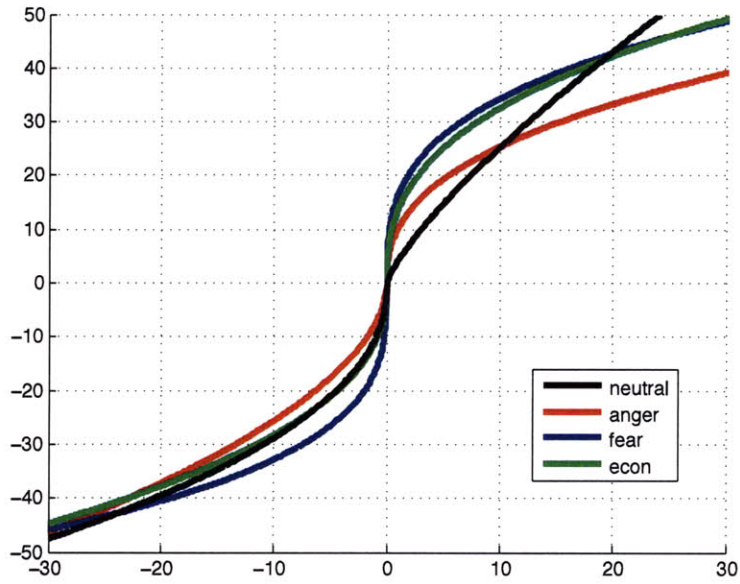
Figure 5-5: Experienced-utility (EU) functions inferred from self-reported EUs in each emotion condition
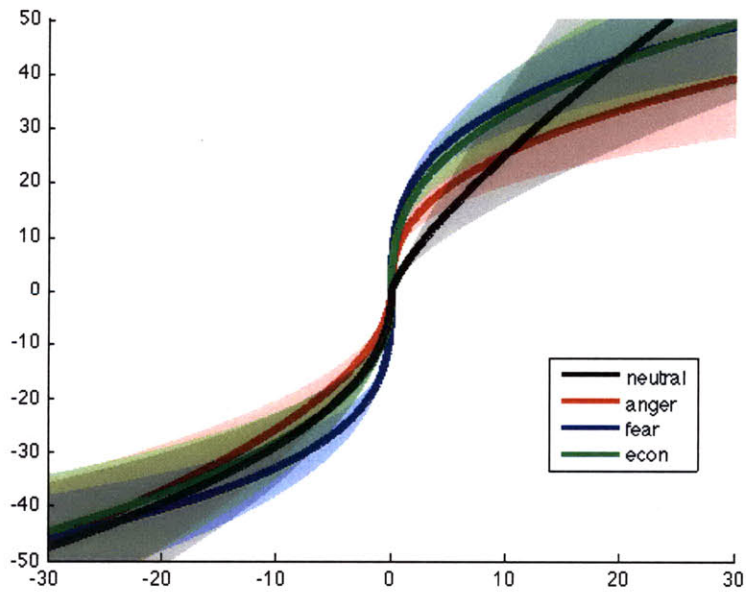
Figure 5-6: Experienced-utility (EU) functions showing the standard error (SE) of the estimated parameter means

(a) Neutral vs. Anger (b) Neutral vs. Fear

(c) Neutral vs. Econ Fear (d) Anger vs. Fear
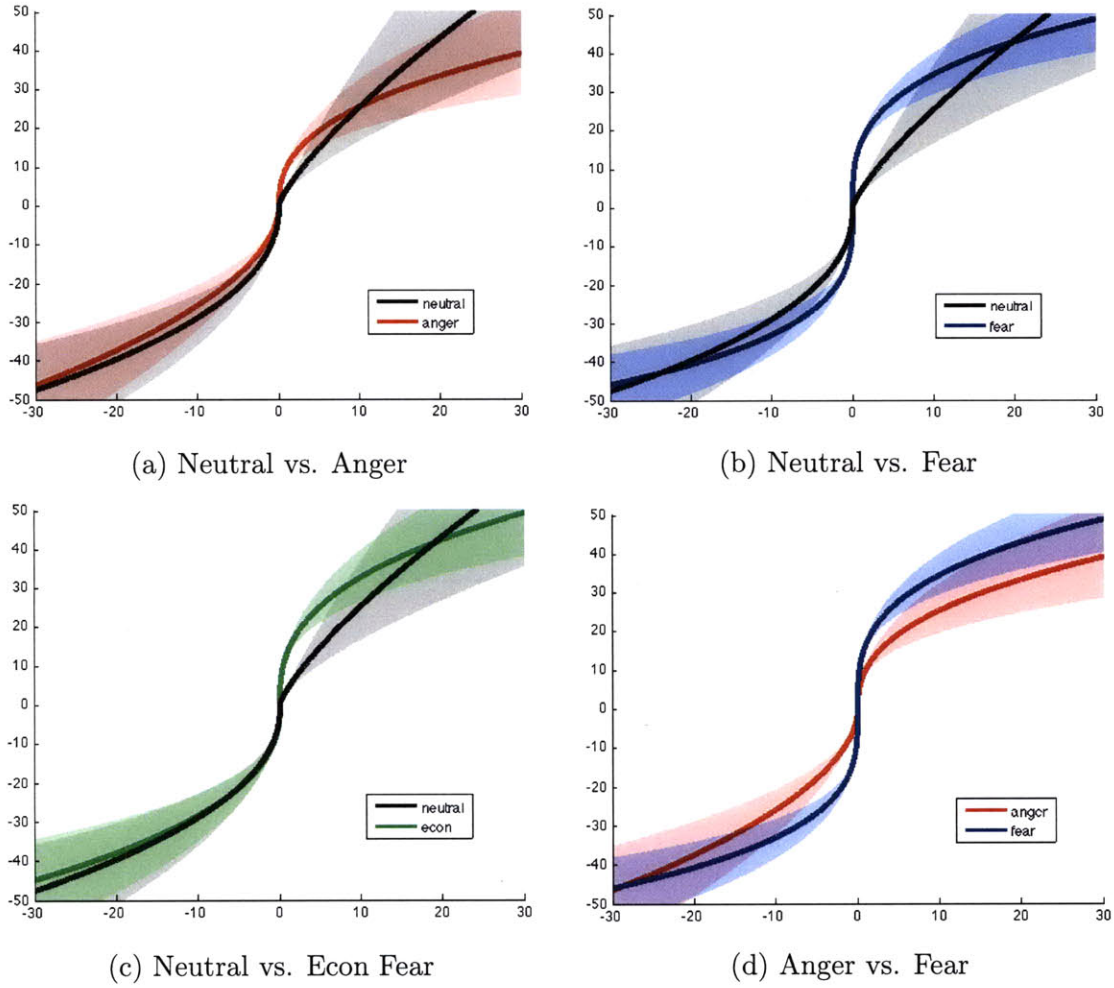
Figure 5-7: Comparisons between two experienced-utility (EU) functions showing the standard error (SE) of the estimated parameter means

more detail.

The self-reported experienced utility $v(t)$ on trial t during a task depends on the obtained outcome $x(t)$, the reference point x_{ref} and the shape parameters $(a, b, \lambda_G, \lambda_L)$ of the experienced-utility function

$$f_{EU}(x(t)|a, b, \lambda_G, \lambda_L, x_{ref}) = \begin{cases} \lambda_G(x(t) - x_{ref})^a, & x(t) \geq x_{ref} \\ -\lambda_L(x_{ref} - x(t))^b, & x(t) < x_{ref} \end{cases} \qquad (5.1)$$

Also, note that NLMEFIT returns the estimated random effects $\delta_{1,j}$ and $\delta_{2,j}$ for each

132

	Neutral	Anger	Fear	Economic Fear
Fixed effects	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)
a	0.77 (0.08)	0.39 (0.05)	0.32 (0.03)	0.37 (0.04)
b	0.45 (0.04)	0.54 (0.04)	0.30 (0.03)	0.42 (0.04)
λ_G	4.25 (0.95)	10.22 (1.39)	16.47 (1.34)	13.76 (1.48)
λ_L	10.19 (1.31)	7.48 (0.98)	16.40 (1.36)	10.89 (1.30)
$\bar{x}_{ref1}$	-0.47 (2.03)	3.94 (2.49)	1.79 (0.10)	2.07 (2.01)
$\bar{x}_{ref2}$	2.09 (1.91)	3.21 (2.36)	2.17 (0.11)	2.09 (1.63)
Random effects	Variance	Variance	Variance	Variance
$\delta_{1,j}$	80.33	122.32	51.04	78.73
$\delta_{2,j}$	73.09	112.09	74.33	54.58
(-) Log Likelihood	4419.1	4449.0	4458.2	4557.6
MSE	231.4	240.7	226.6	290.8
AIC	8856.2	8916.0	8934.5	9133.2
BIC	8900.7	8960.6	8979.1	9177.8

Table 5.2: The EU mixed-effect model parameter estimation for each emotion condition

respondent $j = 1$ to 21 as well as the variances $\text{var}(\delta_{1,j})$ and $\text{var}(\delta_{2,j})$. Thus, the reference points each respondent employed during two tasks can be inferred: $x_{ref1,j} = \bar{x}_{ref1} + \delta_{1,j}$ and $x_{ref2,j} = \bar{x}_{ref2} + \delta_{2,j}$. Two reference points inferred from the EU mixed-effect model will be applied to determine respondents' actually used EU frames (gain or loss frame for experience) for each task and to reveal the influences of the EU frames on their choices and subjective discriminability. Note that the EU frame for each task may be different from the (given) frame for that task which was set by the experimenter. The t-tests $t_{A-B} = (\text{Mean}_A - \text{Mean}_B)/\sqrt{\text{SE}_A^2 + \text{SE}_B^2}$ comparing each parameter between any two emotion conditions revealed the following:

- t-tests comparing the parameter a (curvature in the face of gains): Neutral (0.77) > Anger (0.39), Economic Fear (0.37), Fear (0.32) $ps < .001$

- t-tests comparing the parameter b (curvature in the face of losses): Anger (0.54) > Economic Fear (0.42) > Fear (0.30) $ps < .05$

  Neutral (0.45) > Fear (0.30) $p < .001$

- t-tests comparing the parameter λ_G (sensitivity in the face of gains): Fear (16.47) >

Anger (10.22) > Neutral (4.25) $ps < .001$

Economic Fear (13.76) > Anger (10.22) $p < .05$

- t-tests comparing the parameter λ_L (sensitivity in the face of losses): Fear (16.40) > Economic Fear (10.89), Neutral (10.19) > Anger (7.48) $ps < .05$

First, the parameter a (curvature in the face of gains) is significantly greater (closer to linear) in the neutral condition than in other negative-emotion (anger, fear, economic fear) conditions, explaining that neutral respondents were more risk seeking after experienced gains (i.e., in the EU gain frame).

Second, the parameter b (curvature in the face of losses) is significantly greater (closer to linear) in the neutral and anger conditions than in the fear condition, explaining that fearful respondents were more risk seeking after experienced losses (i.e., in the EU loss frame).

Third, angry respondents (greater b, smaller λ_L) tended to be less sensitive to relatively small experienced losses, compared to respondents in neutral, fear and economical fear conditions, explaining that angry respondents were more risk averse after experienced losses (i.e., in the EU loss frame).

Fourth, fearful and economically fearful respondents (smaller b, greater λ_L) tended to be more sensitive to relatively small experienced gains, compared to neutral and angry respondents, explaining that fearful and economically fearful respondents were more risk averse after explained gains (i.e., in the EU gain frame).

Note that angry respondents were less sensitive to both experienced losses and gains, compared with fearful respondents. Also, the employed EU frame as well as the emotion condition influenced respondents' choice behavior under uncertainty.

Appendix A shows the EU function in each emotion condition with self-reported EU data, and the EU function with each respondent's *inferred* EU reference point on Domain 1 and Domain 2.

5.2.2 Comparing different models for the EU data fit

Different models were compared to fit the self-reported EU data in terms of the BIC (Bayesian Information Criterion). Compared to the maximum likelihood criterion, the

BIC criterion takes account of the number of model parameters (i.e., model complexity) as well as the likelihood in order to avoid overfitting. Thus, BIC is a better criterion for the model comparison. BIC is defined as follows:

$$\text{BIC} = -2 \cdot \ln L + k \ln(n)$$

where

L = the maximized value of the likelihood function for the estimated model

k = the number of free parameters to be estimated

n = the number of data points (i.e., sample size)

Considering the parameterized EU function

$$f_{EU}(x(t)|a, b, \lambda_G, \lambda_L, x_{ref}) = \begin{cases} \lambda_G(x(t) - x_{ref})^a, & x(t) \geq x_{ref} \\ -\lambda_L(x_{ref} - x(t))^b, & x(t) < x_{ref} \end{cases}$$

the following models were compared:

- M1. Optimizing $a, b, \lambda_G, \lambda_L, x_{ref1}$ (reference point on Domain 1), x_{ref2} (reference point on Domain 2) assuming all variables are fixed effects (shared across respondents)

- M2. Assuming $a = 1, b = 1$ (Linear model), Optimizing λ_G, λ_L (fixed effects), x_{ref1}, x_{ref2} (random effects varied across respondents)

- M3. Assuming $\lambda_G = 1$ (PT function model), Optimizing a, b, λ_L (fixed effects), x_{ref1}, x_{ref2} (random effects)

- M4. Assuming $x_{ref1} = x_{ref2} = x_{ref}$, Optimizing $a, b, \lambda_G, \lambda_L$ (fixed effects), x_{ref} (random effects)

- M5. Optimizing $a, b, \lambda_G, \lambda_L$ (fixed effects), x_{ref1}, x_{ref2} (random effects)

- M6. Optimizing $a, b, \lambda_G, \lambda_L, x_{ref1}, x_{ref2}$ (all random effects)

It turned out that the data were not sufficiently numerous for being fitted with the M6 model. The M5 model (on which the dissertation focuses) was the best in terms of the BIC.

	Domain 1		Domain 2	
	Gain EU frame	Loss EU frame	Gain EU frame	Loss EU frame
Neutral	1.32 (N=13)	-0.39 (N=6)	1.74 (N=5)	0.81 (N=16)
	$T(17) = 3.95, p < .001$		$T(19) = 1.36,\ NS$	
Anger	2.40 (N=9)	0.23 (N=8)	0.31 (N=8)	1.29 (N=12)
	$T(15) = 3.21, p < .01$		$T(18) = -2.00, p < .05$	
Fear	2.40 (N=13)	-0.74 (N=8)	-0.11 (N=9)	2.04 (N=12)
	$T(19) = 5.20, p < .001$		$T(19) = -4.31, p < .001$	
Economic Fear	1.65 (N=12)	-0.94 (N=7)	-0.74 (N=5)	2.00 (N=15)
	$T(17) = 3.32, p < .01$		$T(19) = -4.52, p < .001$	
Overall	1.94 (N=47)	-0.46 (N=29)	0.30 (N=27)	1.54 (N=55)
	$F(1,75) = 56.7, p < .001$		$T(1,81) = 18.7, p < .001$	

Table 5.3: Discriminability ratio = subjective d' / objective d'

5.2.3 The domain, the EU frame and the emotion condition influence the discriminability ratio d'_{subj}/d'_{obj}

Respondents' objective discriminability values were not significantly different for all pairwise comparisons (across the gain and EU loss frames and across the domain types, and across the emotion condition). Their subjective discriminability values, however, were significantly influenced by the EU frame and the domain type for each emotion condition.

Table 5.3 shows how the EU frame (gain or loss frame) influenced the discriminability ratio ($= d'_{subj}/d'_{obj}$) on each domain for emotion conditions. Note that only the tasks in which both options were selected for at least two trials were included in this discriminability analysis. This is the required condition for estimating the standard deviations of sampled outcomes or experienced utilities.

On Domain 1 (where optimal option = safe option), respondents across all emotion conditions showed greater discriminability ratios leading to more risk-aversion and optimal selections in the EU gain frame than in the EU loss frame (supporting H4-1). Respondents employing the EU gain frame made more risk-averse choices (optimal choices) more often on Domain 1, compared to "rational" decision makers who were assumed to make decisions based on objective outcomes (not based on subjective utilities).

On Domain 2 (where optimal option = risky option), respondents in emotion conditions

136
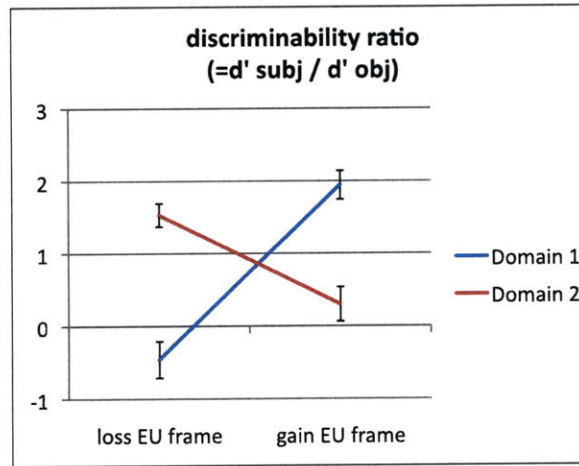
Figure 5-8: The interaction effect of the domain type and the EU frame on the discriminability ratio (pooling over respondents in all four emotion conditions). Note that error bars indicate standard errors of the mean.

(a) Domain 1   (b) Domain 2
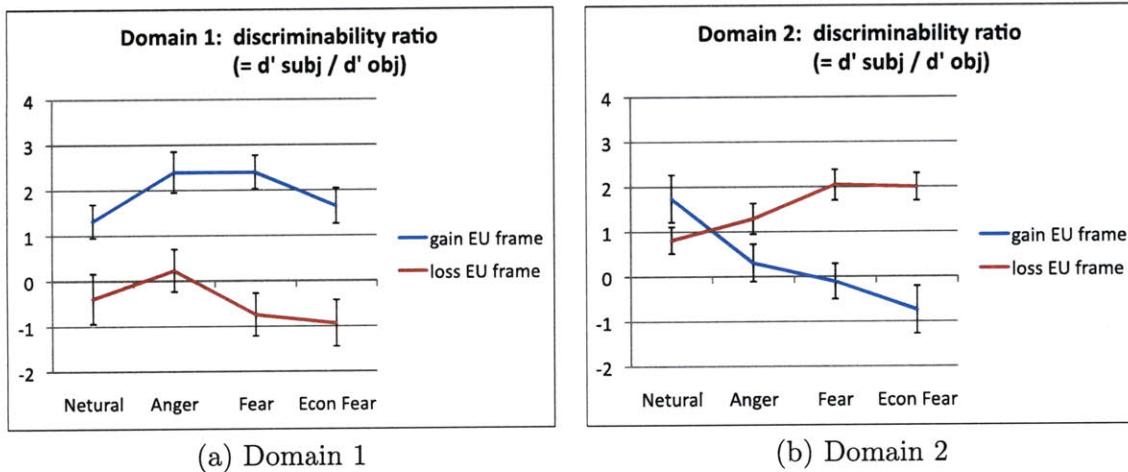
Figure 5-9: The interaction effect of the EU frame and the emotion condition on discriminability ratio for each domain. Note that error bars indicate standard errors of the mean.

except the neutral condition showed greater discriminability ratios (leading to more risk seeking and optimal selections) in the EU loss frame than in the EU gain frame (partly supporting H4-1). Respondents employing the EU loss frame made more risk-seeking choices (optimal choices) more often on Domain 2, compared to "rational" decision makers who were assumed to make decisions based on objective outcomes (not based on subjective utilities). Note that neutral respondents in the EU gain frame showed greater discriminability ratio

than in the EU loss frame on Domain 2, but this was not statistically significant ($T(19) =$ 1.36, NS). This can be associated with the fact that the parameter $a= 0.77$, curvature in the face of gains of the neutral condition was greater than that of other conditions, leading to more risk seeking after experienced gains.

Moreover, on each domain, the emotion condition influenced the discriminability ratio, depending on the EU frame:

On Domain 1, fearful respondents employing the EU gain frame tended to have greater discriminability ratios (or greater risk aversion) than neutral respondents employing the same frame ($M_{neutral}=1.32$ vs. $M_{fear} = 2.40$; $T(24) = -2.45$, $p < .05$).

On Domain 2, neutral or angry respondents employing the EU loss frame tended to have smaller discriminability ratios (or less risk aversion) than fearful or economically fearful respondents employing the same frame (pairwise T-test comparisons, $ps< .05$).

On Domain 2, fearful or economically fearful respondents showed negative discriminability ratios (or more risk aversion) in the EU gain frame, whearas neutral or angry respondents showed positive discriminability ratios (or more risk seeking) in that frame, but these were not statistically significant by pairwise T-test comparisons.

5.2.4 Experience, gender and emotion influence confidence, which influences future prediction and decision

Association between experience-utility and confidence

Respondents' experienced utility $v(t)$ from the obtained outcome on trial t influenced their confidence $q(t)$. For each emotion condition (neutral, anger, fear, economic fear), the average of confidence over trials with relatively higher experienced utility ($v(t) > E_\tau[v(\tau)]$, trials on which the experienced utility was greater than the average experienced-utility over the task) was significantly greater than the average of confidence over trials with relatively lower experienced utility ($v(t) < E_\tau[v(\tau)]$) ($ps< .05$, Table 5.4). This means that when respondents had a relatively good experience from the current trial, they tended to get more confident in their next trial, compared to when they had a relatively bad experience.

There was a significantly high correlation between the moving-averaged experienced-utility and the moving-averaged confidence, $corr = 0.77, p < .001$ using equally-weighted

138

10-point moving-average filter, and $corr = 0.70, p < .001$ using equally-weighted 5-point moving-average filter, pooling over respondents across emotion conditions (neutral, anger, fear, economic fear).

	Averaged Confidence	
	Lower EU	Higher EU
Neutral	-1.37	10.39
	$T(41) = 6.31, p < .001$	
Anger	-9.80	3.26
	$T(41) = 7.75, p < .001$	
Fear	-7.23	7.06
	$T(41) = 5.73, p < .001$	
Economic Fear	-12.14	-1.47
	$T(41) = 5.74, p < .001$	
Overall	-7.64	4.81
	$T(167) = 12.50, p < .001$	

Table 5.4: Comparing the means of confidence for lower-EU trials and higher-EU trials on each emotion condition

	PU difference	
	Low Confidence	High Confidence
Neutral	17.58	31.71
	$T(36) = 3.50, p < .001$	
Anger	19.33	34.28
	$T(31) = 3.25, p < .005$	
Fear	19.97	35.27
	$T(29) = 3.53, p < .001$	
Economic Fear	27.56	36.83
	$T(31) = 2.42, p < .05$	
Overall	21.12	34.43
	$T(130) = 6.42, p < .001$	

Table 5.5: Comparing the means of absolute PU difference for lower-EU trials and higher-EU trials on each emotion condition

Association between confidence and predicted-utility difference

Respondents' confidence $q(t)$ influenced the absolute difference of their predicted utilities $|u(t|1) - u(t|2)|$ before their next trial. For respondents in each emotion condition (neutral, anger, fear, economic fear), the average of absolute PU differences over trials with high confidence ($50 \geq q(t) > 0$) was significantly greater than the average of absolute PU differences over trials with low confidence ($0 > q(t) \geq -50$) ($ps < .05$, Table 5.5). This means that

139

	Pr (option = exploitative choice)	
	Lower PU difference	Higher PU difference
Neutral	59%	75%
	$T(41) = 4.37, p < .001$	
Anger	61%	78%
	$T(40) = 4.01, p < .001$	
Fear	52%	74%
	$T(40) = 5.84, p < .001$	
Economic Fear	56%	77%
	$T(41) = 5.26, p < .001$	
Overall	57%	76%
	$T(165) = 9.75, p < .001$	

Table 5.6: Comparing the means of exploitative-choice probability for lower PU-difference trials and higher PU-difference trials on each emotion condition

when respondents were in a high confidence state, they tended to predict that one option was much better than the other, compared to when they were in a low confidence state.

Association between predict-utility difference and exploitative-choice probability

Respondents tended to select the exploitative choice (= option with the current-estimated greatest predicted utility) more often as they had higher PU difference of two options ($|u(t|1) - u(t|2)|$). Table 5.6 compares the probabilities of selecting the exploitative choice over higher PU difference trials (i.e., trials on which the PU difference is greater than the average PU difference over all trials) and over lower PU difference trials (i.e., trials on which the PU difference is smaller than the average PU difference over all trials). Higher PU difference increased the exploitative-choice probability (resulting in less randomized choice behavior), whereas lower PU difference increased the exploratory-choice probability (resulting in more randomized choice behavior). This supports the hypothesis on how the confidence state influences the sensitivity of the predicted-utility function, which regulates the trade-offs between exploitation and exploration.

Gender differences in confidence and predict-utility difference

Tables 5.7 and 5.8, respectively, show the associations between experienced-utility and confidence for male and female respondents. Tables 5.9 and 5.10, respectively, show the

140

associations between confidence and predict-utility difference for male and female respondents. In Table 5.8, note that fearful female respondents showed negative (very low) average confidence both over lower-EU trials (M=-19.41) and over higher-EU trials (M=-9.55). In other words, fearful female respondents tended to feel less confident even with previous (relatively) higher experienced utility ($v(t) > E_\tau[v(\tau)]$), possibly due to the influence of the existing incidental fearful emotion. Table 5.13 shows the comparisons between male and female respondents in terms of experienced-utility and confidence. In the fear condition, female respondents showed significantly lower average confidence than male respondents for both cases of previous lower EU (M_{male}=-2.36 vs. M_{female}=-19.41, $T(40) = 2.23$, $p < .05$) and of previous higher EU (M_{male}=13.70 vs. M_{female}=-9.55, $T(40) = 2.89$, $p < .01$).

	Average Confidence	
	Lower EU	Higher EU
Neutral	0.92	13.70
	$T(21) = 5.29, p < .001$	
Anger	-7.84	3.01
	$T(25) = 6.19, p < .001$	
Fear	-2.36	13.70
	$T(29) = 5.15, p < .001$	
Economic Fear	-12.17	-0.81
	$T(19) = 3.41, p < .005$	
Overall	-5.08	7.91
	$T(97) = 9.48, p < .001$	

Table 5.7: Male respondents: Comparing the means of confidence for lower-EU trials and higher-EU trials on each emotion condition

	Average Confidence	
	Lower EU	Higher EU
Neutral	-3.89	6.75
	$T(19) = 3.65, p < .005$	
Anger	-12.99	3.66
	$T(15) = 5.10, p < .001$	
Fear	-19.41	-9.55
	$T(11) = 2.59, p < .05$	
Economic Fear	-12.11	-2.07
	$T(21) = 5.22, p < .001$	
Overall	-11.21	0.48
	$T(69) = 8.15, p < .001$	

Table 5.8: Female respondents: Comparing the means of confidence for lower-EU trials and higher-EU trials on each emotion condition

141

	Average PU difference	
	Low Confidence	High Confidence
Neutral	17.87	37.98
	$T(17) = 3.48, p < .005$	
Anger	24.12	30.46
	$T(20) = 2.50, p < .05$	
Fear	27.88	45.46
	$T(20) = 3.45, p < .005$	
Economic Fear	24.97	32.71
	$T(15) = 2.32, p < .05$	
Overall	23.86	36.86
	$T(75) = 5.74, p < .001$	

Table 5.9: Male respondents: Comparing the means of absolute PU difference for low-confidence trials and high-confidence trials on each emotion condition

	Average PU difference	
	Low Confidence	High Confidence
Neutral	19.07	24.60
	$T(18) = 1.43, p = .09$	
Anger	17.05	37.40
	$T(10) = 2.45, p < .05$	
Fear	11.69	15.42
	$T(8) = 1.57, p = .08$	
Economic Fear	32.65	40.48
	$T(15) = 1.39, p = .10$	
Overall	21.41	30.28
	$T(54) = 3.21, p < .005$	

Table 5.10: Female respondents: Comparing the means of absolute PU difference for low-confidence trials and high-confidence trials on each emotion condition

On low-confidence trials, fearful female respondents reported significantly lower average PU difference than neutral female respondents ($M_{neutral}=19.07$ vs. $M_{fear}=11.69$, $T(29)=2.36$, $p < .05$) (Table 5.10). On high-confidence trials, fearful female respondents reported lower average PU difference than neutral female respondents on average, but this was not statistically significant ($M_{neutral}=24.60$ vs. $M_{fear}=15.42$, $T(27)=1.29$, $p < .10$).

Since lower confidence led to lower PU difference (Tables 5.9 and 5.10) and lower PU difference led to more randomized choice behavior (Tables 5.11 and 5.12), fearful female respondents tended to have more exploratory trials (i.e., more randomized choices) than fearful male respondents. In other words, fearful female respondents selected fewer exploitative choices (i.e., fewer selections of the option with the current-estimated greatest

	Pr (option = exploitative choice)	
	Lower PU difference	Higher PU difference
Neutral	63%	79%
	$T(21) = 2.70, p < .01$	
Anger	64%	77%
	$T(25) = 2.98, p < .01$	
Fear	55%	79%
	$T(29) = 5.41, p < .001$	
Economic Fear	56%	75%
	$T(19) = 3.42, p < .01$	
Overall	59%	78%
	$T(97) = 7.33, p < .001$	

Table 5.11: Male respondents: Comparing the means of exploitative-choice probability for lower PU-difference trials and higher PU-difference trials on each emotion condition

	Pr (option = exploitative choice)	
	Lower PU difference	Higher PU difference
Neutral	55%	72%
	$T(19) = 3.68, p < .005$	
Anger	57%	78%
	$T(14) = 2.66, p < .01$	
Fear	44%	60%
	$T(10) = 2.29, p < .05$	
Economic Fear	55%	79%
	$T(21) = 3.93, p < .001$	
Overall	54%	73%
	$T(67) = 6.39, p < .001$	

Table 5.12: Female respondents: Comparing the means of exploitative-choice probability for lower PU-difference trials and higher PU-difference trials on each emotion condition

	Average Confidence (Lower EU)		Average Confidence (Higher EU)	
	Male	Female	Male	Female
Neutral	0.92	-3.89	13.70	6.75
	$T(40) = 0.716, NS$		$T(40) = 1.07, NS$	
Anger	-7.84	-12.99	3.01	3.66
	$T(40) = 0.78, NS$		$T(40) = -0.08, NS$	
Fear	-2.36	-19.41	13.70	-9.55
	$T(40) = 2.23, p < .05$		$T(40) = 2.89, p < .01$	
Economic Fear	-12.17	-12.11	-0.81	-2.07
	$T(40) = -0.01, NS$		$T(40) = 0.15, NS$	
Overall	-5.08	-11.21	7.91	0.48
	$T(153) = 1.09, p = .14$		$T(142) = 1.65, p = .05$	

Table 5.13: Comparisons between male and female respondents on each emotion condition: Confidence for lower-EU trials and higher-EU trials

143

	Average PU difference (Low Confidence)		Average PU difference (High Confidence)	
	Male	Female	Male	Female
Neutral	17.87	19.07	37.98	24.60
	$T(37) = -0.76$, NS		$T(38) = 2.38$, $p < .05$	
Anger	24.12	17.05	30.46	37.40
	$T(38) = 1.38$, $p = .09$		$T(32) = -0.64$, NS	
Fear	27.88	11.69	45.46	15.42
	$T(35) = 2.26$, $p < .05$		$T(33) = 3.11$, $p < .005$	
Economic Fear	24.97	32.65	32.71	40.48
	$T(37) = -0.26$, NS		$T(33) = -1.30$, $p = .10$	
Overall	23.86	21.41	36.86	30.28
	$T(166) = 1.77$, $p < .05$		$T(166) = 1.92$, $p < .05$	

Table 5.14: Comparison between male and female respondents on each emotion condition: Absolute PU difference for lower-EU trials and higher-EU trials

	Pr (option = exploitative choice) (Lower PU difference)		Pr (option = exploitative choice) (Higher PU difference)	
	Male	Female	Male	Female
Neutral	63%	55%	79%	72%
	$T(40) = 1.05$, NS		$T(40) = 1.18$, NS	
Anger	64%	57%	77%	78%
	$T(39) = 0.83$, NS		$T(39) = -0.05$, NS	
Fear	55%	44%	79%	60%
	$T(40) = 1.57$, $p = .06$		$T(39) = 2.12$, $p < .05$	
Economic Fear	56%	55%	75%	79%
	$T(40) = 0.13$, NS		$T(40) = -0.44$, NS	
Overall	59%	54%	78%	73%
	$T(165) = 1.51$, $p = .07$		$T(164) = 1.25$, NS	

Table 5.15: Comparisons between male and female respondents on each emotion condition: Exploitative-choice probability for lower PU-difference trials and higher PU-difference trials

predicted utility) than fearful male respondents ($M_{male} = 71.9\%$ vs. $M_{female} = 54.7\%$ for the exploitative-choice probability over all trials, $T(40) = 2.70$, $p < .01$).

The greater exploitative-choice probability of male respondents in the fear condition made them more often select the optimal option (i.e., the safe option on Domain 1, the risky option on Domain 2) on average than female respondents, but this was not statistically significant ($M_{male} = 65.5\%$ vs. $M_{female} = 62.5\%$ for the optimal-choice probability over all trials, $T(40) = 0.46$, NS) (Table 5.21). In addition, note that 68.5% of the exploitative choices of male respondents in the fear condition were optimal option selections, whereas 62.5% of the exploitative choices of female respondents in that condition were optimal option selections (Table 5.19).

144

Figure 5-10 presents the proposed model on the interactions of gender, emotion, experience, confidence, prediction and exploitative/exploratory decisions. Gender and emotion (e.g., incidental fear) may influence confidence. Good (bad) experience leads to high (low) confidence. Also, then, high (low) confidence leads to high (low) predicted-utility difference and more exploitative (exploratory) choice behavior.
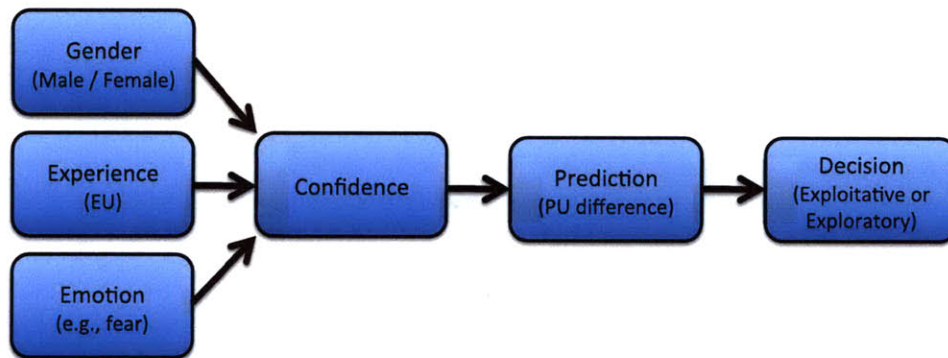
Figure 5-10: Gender and emotion may influence confidence. Good (bad) experience leads to high (low) confidence. Also, then, high (low) confidence leads to high (low) predicted-utility difference and more exploitative (exploratory) choice behavior

	Average EU	
	Male	Female
Neutral	-1.4	-4.7
	$T(40) = 0.51, NS$	
Anger	-2.6	-3.1
	$T(40) = 0.07, NS$	
Fear	5.7	-11.9
	$T(40) = 2.23, p < .05$	
Economic Fear	-2.6	3.5
	$T(40) = -0.95, NS$	
Overall	0.2	-3.0
	$T(166) = 0.93, NS$	

Table 5.16: Comparisons between male and female respondents on each emotion condition: Average EU over trials

	Average Confidence	
	Male	Female
Neutral	8.0	0.5
	$T(40) = 1.13, NS$	
Anger	-2.1	-2.8
	$T(40) = 0.09, NS$	
Fear	7.9	-15.7
	$T(40) = 3.01, p < .005$	
Economic Fear	-6.2	-5.8
	$T(40) = -0.06, NS$	
Overall	2.4	-5.0
	$T(166) = 1.98, p < .05$	

Table 5.17: Comparisons between male and female respondents on each emotion condition: Avearge confidence over trials

	Pr (option = exploitative choice)	
	Male	Female
Neutral	75.4%	65.3%
	$T(40) = 1.68, p = .05$	
Anger	72.4%	65.6%
	$T(40) = 0.91, NS$	
Fear	71.9%	54.7%
	$T(40) = 2.70, p < .01$	
Economic Fear	68.5%	73.3%
	$T(40) = -0.70, NS$	
Overall	72.1%	66.1%
	$T(166) = 1.82, p < .05$	

Table 5.18: Comparisons between male and female respondents on each emotion condition: Exploitative-choice probability over all trials

	Pr (option = optimal | exploitative choice)	
	Male	Female
Neutral	71.1%	70.0%
	$T(40) = 0.14, NS$	
Anger	65.6%	77.5%
	$T(39) = -1.35, NS$	
Fear	68.5%	62.5%
	$T(40) = 0.06, NS$	
Economic Fear	69.2%	58.5%
	$T(40) = 1.05, NS$	
Overall	68.4%	66.6%
	$T(165) = 0.40, NS$	

Table 5.19: Comparisons between male and female respondents on each emotion condition: Optimal-choice probability over exploitative trials

	Pr (option = optimal | exploratory choice)	
	Male	Female
Neutral	52.1%	53.0%
	$T(36) = -0.08, NS$	
Anger	42.6%	60.0%
	$T(31) = -1.39, NS$	
Fear	47.5%	47.7%
	$T(39) = -0.02, NS$	
Economic Fear	45.3%	44.9%
	$T(38) = 0.03, NS$	
Overall	46.7%	50.6%
	$T(150) = -0.71, NS$	

Table 5.20: Comparisons between male and female respondents on each emotion condition: Optimal-choice probability over exploratory trials

	Pr (option = optimal choice)	
	Male	Female
Neutral	65.9%	62.8%
	$T(40) = 0.55, NS$	
Anger	65.5%	66.0%
	$T(40) = -0.08, NS$	
Fear	65.5%	62.5%
	$T(40) = 0.46, NS$	
Economic Fear	68.0%	56.2%
	$T(40) = 1.61, p = .06$	
Overall	66.1%	61.4%
	$T(166) = 1.49, p = .07$	

Table 5.21: Comparisons between male and female respondents on each emotion condition: Optimal-choice probability over all trials

5.3 The analysis based on the PU frame

The PU frame is the frame the decision maker actually employed to estimate predicted utility and may or may not be the same as the given frame. The PU frame can be inferred from the data-fitting analysis of respondents' self-reported confidences and predictions during the decision task. The PU frame-based analysis presents the interaction effects of framing and emotion on human prediction and decisions.

147

5.3.1 Emotion conditions influence the PU frame and the shape of the predicted-utility function

Respondents tended to use the bet money as their PU reference point (i.e., the reference point of their PU function estimated by the PU data fitting) in the (given) gain frame on Domain 1, in the loss frame on Domain 2, and in the gain frame on Domain 2. In other words, except in the loss frame on Domain 1, respondents in all emotion conditions tended to involve the PU gain frame when the bet money is smaller than the average outcome of the safe option, and involve the EU loss frame when the bet money is greater than the average outcome of the safe option.

Note that for the loss frame on Domain 1, respondents first received the bet money of $14 or $20 and then lost some money after selecting an option on each trial. Also, the average outcome of the safe option was a loss of $18.

For the given loss frame on Domain 1, respondents in the neutral condition tended to use the bet money as their PU reference point. In other words, neutral-emotion respondents involved the PU loss frame when they received the bet money of $14 (i.e., self-reporting displeasure for the expected loss of $18 from the safe option), but involved the PU gain frame when they received the bet money of $20 (i.e., self-reporting pleasure for the expected loss of $18 from the safe option). That is, they tended to be very rational in light of the net outcome. Yet, *respondents in the fear condition tended to employ the PU loss frame* self-reporting displeasure for the expected loss of $18 from the safe option regardless of receiving $14 or $20 initially (Fisher's exact test, $p < .01$); their PU reference point tended to be greater than the average outcome of the safe option (i.e., the PU loss frame), so they view the expected outcomes from the safe option as losses regardless of their bet-money condition.

Now the following analysis presents the inferred predicted-utility function from self-reported PUs in each emotion condition. Appendix C describes the PU data-fit method in more detail.

The predicted-utility sample $y(\tau)$ corresponding to outcome $x(\tau)$ ($\tau = 1$ to t) is computed using the predicted utility function:
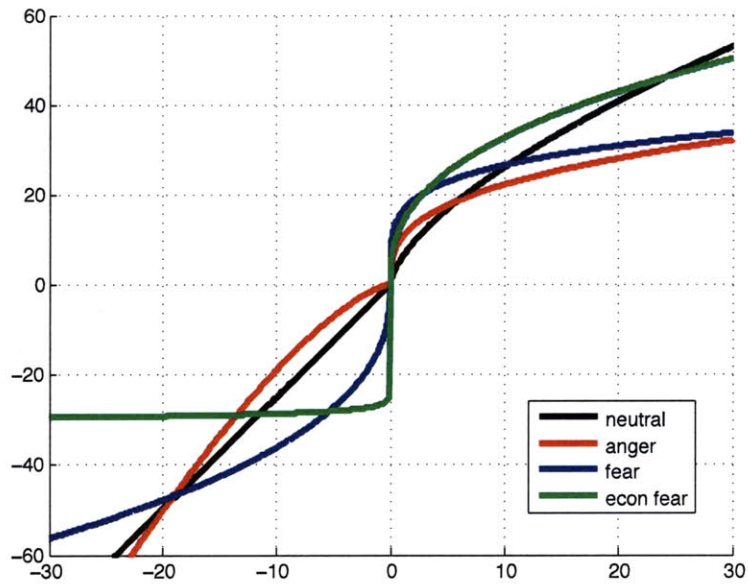
Figure 5-11: Predicted-utility (PU) functions inferred from self-reported PUs in each emotion condition, drawing the functions when $e(t)=0.5$ (neutral confidence)
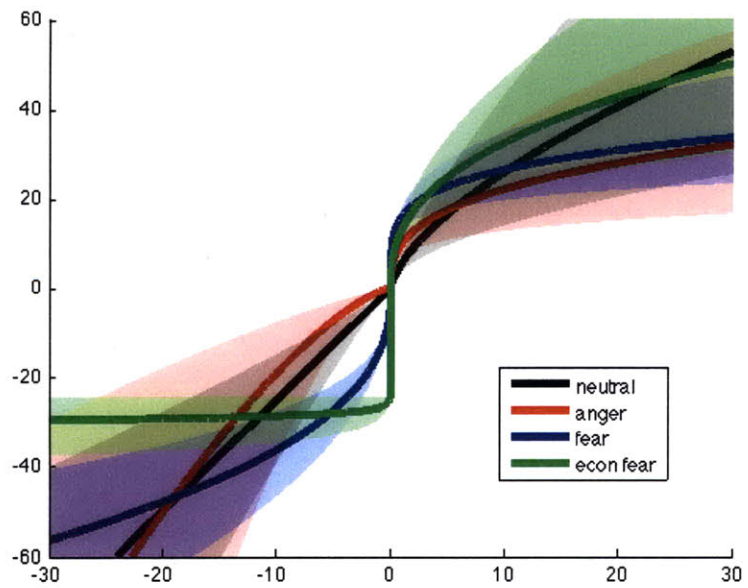
Figure 5-12: Predicted-utility (PU) functions showing the standard error (SE) of the estimated parameter means (drawing the functions when $e(t)=0.5$ (neutral confidence))

(a) Neutral vs. Anger

(b) Neutral vs. Fear

(c) Neutral vs. Econ Fear

(d) Anger vs. Fear
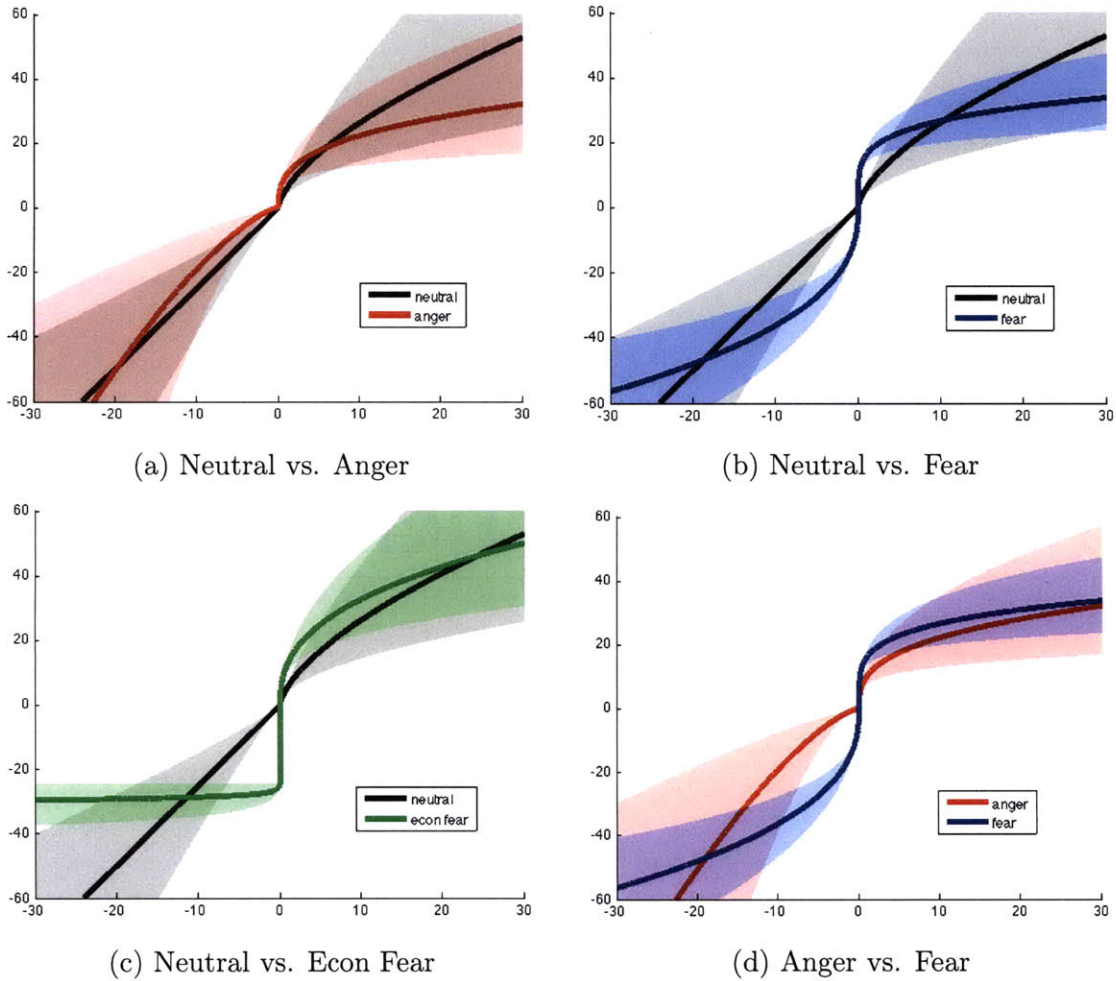
Figure 5-13: Comparisons between two predicted-utility (PU) functions showing the standard error (SE) of the estimated parameter means (drawing the functions when $e(t)=0.5$ (neutral confidence))

$$f_{PU}(x(\tau)|a, b, \lambda_G(\tau), \lambda_L(\tau), x_{ref}) = y(\tau) = \begin{cases} \lambda_G(\tau)\,(x(\tau) - x_{ref})^a, & x(\tau) \geq x_{ref} \\ -\lambda_L(\tau)\,(x_{ref} - x(\tau))^b, & x(\tau) < x_{ref} \end{cases}$$

Note that sensitivity parameters $\lambda_G(t)$ and $\lambda_L(t)$ depend on the self-reported confidence measurement $q(t)$. Thus, $y(\tau)$ depends on the reference point x_{ref} and the shape parameters $(a, b, \lambda_G(t), \lambda_L(t))$ of the predicted utility function.

Sensitivity parameters $\lambda_G(t)$ and $\lambda_L(t)$ are functions of the confidence state variable $e(t)$

(a) Neutral

(b) Anger

(c) Fear

(d) Economic Fear
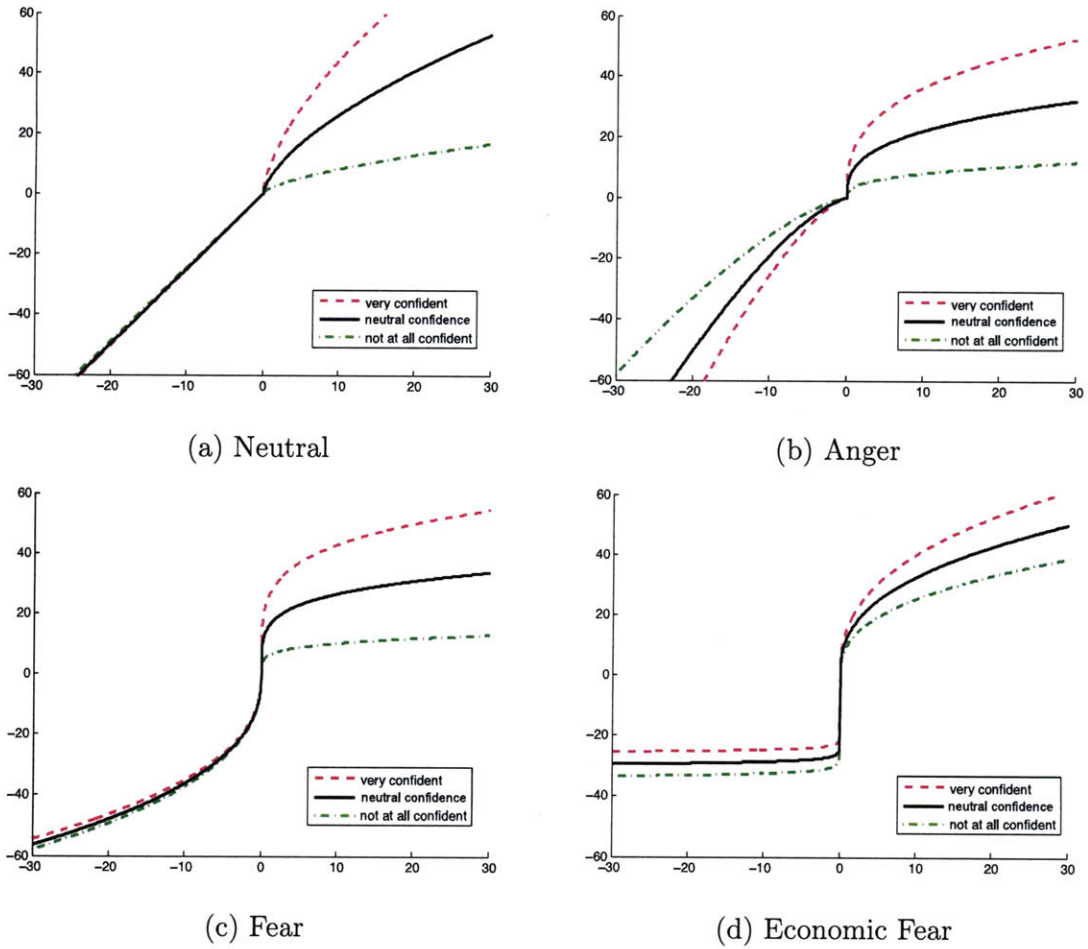
Figure 5-14: Predicted utility (PU) function changing with confidence: the magenta line when $e(t) = 1$ (very confident), the black line when $e(t) = 0.5$ (neutral confidence), and the green line when $e(t) = 0$ (not at all confident)

ranging from 0 (= not at all confident) to 1 (= very confident):

$$e(t) = \frac{1}{1 + \exp(-\kappa q(t))} \text{where} \kappa = 0.1$$

$$\lambda_G(t) = \lambda_G(e(t)) = \lambda_{Gbase} + \lambda_{Gslope}(2\,e(t) - 1)$$
$$\lambda_L(t) = \lambda_L(e(t)) = \lambda_{Lbase} + \lambda_{Lslope}(2\,e(t) - 1)$$

When $e(t) = 0.5$ (neutral confidence), $\lambda_G(t) = \lambda_{Gbase}$ and $\lambda_L(t) = \lambda_{Lbase}$.

When $e(t) = 0$ (not at all confident), $\lambda_G(t) = \lambda_{Gbase} - \lambda_{Gslope}$ and $\lambda_L(t) = \lambda_{Lbase} - \lambda_{Lslope}$.

151

	Neutral	Anger	Fear	Economic Fear
Fixed effects	Mean (SE)	Mean (SE)	Mean (SE)	Mean (SE)
a	0.64 (0.11)	0.34 (0.10)	0.22 (0.06)	0.39 (0.08)
b	0.99 (0.09)	1.40 (0.14)	0.40 (0.05)	0.03 (0.04)
λ_{Gbase}	5.92 (1.80)	10.16 (2.70)	16.06 (2.35)	13.32 (2.70)
λ_{Lbase}	2.54 (0.71)	0.76 (0.35)	14.37 (2.06)	26.87 (2.52)
λ_{Gslope}	4.07 (1.28)	6.44 (1.82)	9.92 (1.65)	3.02 (0.95)
λ_{Lslope}	0.04 (0.09)	0.26 (0.11)	-0.43 (0.55)	-3.68 (1.48)
$\bar{x}_{ref1}$	1.63 (1.59)	2.01 (1.84)	2.50 (0.87)	0.66 (0.49)
$\bar{x}_{ref2}$	2.60 (1.27)	2.20 (1.55)	1.67 (0.03)	-0.85 (0.02)
Random effects	Variance	Variance	Variance	Variance
$\delta_{1,j}$	42.51	49.57	38.28	13.79
$\delta_{2,j}$	78.57	68.87	35.03	24.08
(-) Log Likelihood	9166.3	9097	9291	9662.2
MSE	338.5	318.6	375.8	538.0
AIC	18355	18217	18605	19346
BIC	18417	18279	18667	19409

Table 5.22: The PU mixed-effect model parameter estimation for each emotion condition

When $e(t) = 1$ (very confident), $\lambda_G(t) = \lambda_{Gbase} + \lambda_{Gslope}$ and $\lambda_L(t) = \lambda_{Lbase} + \lambda_{Lslope}$. Also, $\lambda_G(t)$ and $\lambda_L(t)$ should be always positive. Note that λ_{Gbase} and λ_{Lbase} should be positive values, but that λ_{Gslope} and λ_{Lslope} may be either positive or negative.

NLMEFIT returns the estimated random effects $\delta_{1,j}$ and $\delta_{2,j}$ for each respondent $j = 1$ to 21 as well as the variances $var(\delta_{1,j})$ and $var(\delta_{2,j})$. Thus, the reference points each respondent employed during two tasks can be inferred: $x_{ref1,j} = \bar{x}_{ref1} + \delta_{1,j}$ and $x_{ref2,j} = \bar{x}_{ref2} + \delta_{2,j}$.

Two reference points inferred from the PU mixed-effect model will be applied to determine respondents' actually used PU frames (gain or loss frame for prediction) for each task and to reveal the influences of the PU frames on their risk attitudes and choices. Note that the PU frame for each task may be different from the EU frame or the frame given by the experimenter for that task.

The t-tests $(t_{A-B} = (\text{Mean}_A - \text{Mean}_B)/\sqrt{\text{SE}_A^2 + \text{SE}_B^2})$ comparing each parameter between any two emotion conditions revealed the following:

- t-tests comparing the parameter a (curvature in the face of gains):

152

Neutral (0.64) > Economic Fear (0.39), Anger (0.34), Fear (0.22) $ps < .05$

- t-tests comparing the parameter b (curvature in the face of losses):

  Anger (1.40) > Neutral (0.99) > Fear (0.40) > Economic Fear (0.03) $ps < .01$

- t-tests comparing the parameter λ_{Gbase} (sensitivity base in the face of gains):

  Fear (16.06), Economic Fear (13.32) > Neutral (5.92) $ps < .05$

  Fear (16.06) > Anger (10.16) $p < .05$

- t-tests comparing the parameter λ_{Lbase} (sensitivity base in the face of losses):

  Economic Fear (26.87) > Fear (14.37) > Neutral (2.54) > Anger (0.76) $ps < .05$

- t-tests comparing the parameter λ_{Gslope} (sensitivity slope in the face of gains):

  Fear (9.92) > Neutral (4.07) > Economic Fear (3.02) $ps < .01$

  Anger (6.44) > Economic Fear (3.02) $p < .05$

- t-tests comparing the parameter λ_{Lslope} (sensitivity slope in the face of losses):

  Anger (0.26), Neutral (0.04), Fear (-0.43) > Economic Fear (-3.68) $ps < .05$

- t-tests comparing $\lambda_G(e(t) = 0) = \lambda_{Gbase} - \lambda_{Gslope}$ (not at all confident) with $\lambda_G(e(t) = 1) = \lambda_{Gbase} + \lambda_{Gslope}$ (very confident) in the PU gain frame:

  Neutral (1.85 vs. 9.99) $p < .01$

  Anger (3.72 vs. 16.60) $p < .01$

  Fear (6.14 vs. 25.98) $p < .01$

  Economic Fear (10.30 vs. 16.34) $p = .07$

- t-tests comparing $\lambda_L(e(t) = 0) = \lambda_{Lbase} - \lambda_{Lslope}$ (not at all confident) with $\lambda_L(e(t) = 1) = \lambda_{Lbase} + \lambda_{Lslope}$ (very confident) in the PU loss frame:

  Economic Fear (30.55 vs. 23.19) $p < .05$

  Note: $\lambda_{Lslope}(= -3.68) < 0$ for the economic fear condition

153

First, the parameter a (curvature in the face of gains) was significantly greater (closer to linear) in the neutral condition than in the other negative-emotion (anger, fear, economic fear) conditions, explaining that neutral respondents were more risk seeking in face of potential gains (i.e., in the PU gain frame).

Second, the parameter b (curvature in the face of losses) was significantly greater in the anger condition ($= 1.40 > 1$) than in the dition ($= 0.99$), explaining that angry respondents were more risk averse in face of potential losses (i.e., in the PU loss frame).

Third, the parameter b was significantly smaller in the fear ($= 0.4$) and economic fear ($= 0.03$) conditions than in the neutral condition ($= 0.99$), explaining that fearful and economically fearful respondents were more risk seeking in face of potential losses (i.e., in the PU loss frame).

Fourth, t-tests comparing $\lambda_G(e(t) = 0) = \lambda_{Gbase} - \lambda_{Gslope}$ (not at all confident) with $\lambda_G(e(t) = 1) = \lambda_{Gbase} + \lambda_{Gslope}$ (very confident) in the PU gain frame revealed that respondents' confidence significantly influenced their sensitivity to predicted gains. In all emotion conditions, when respondents were very confident in the task, they tended to be more sensitive to potential gains. But with lower confidence, they were less sensitive to potential gains.

Note that the higher sensitivity to potential gains under higher confidence is very likely to facilitate respondents' optimal decision making under uncertainty. Also, this result is consistent with our hypothesis H5.

Fifth, t-tests comparing $\lambda_L(e(t) = 0) = \lambda_{Lbase} - \lambda_{Lslope}$ (not at all confident) with $\lambda_L(e(t) = 1) = \lambda_{Lbase} + \lambda_{Lslope}$ (very confident) in the loss PU frame revealed that respondents' confidence did not significantly influence their sensitivity to predicted losses, except in the economic fear condition. Interestingly, in the economic fear condition, $\lambda_{Lslope} = -3.68$ was negative; thus, higher confidence reduced the sensitivity to predicted losses. The fear condition also had a negative $\lambda_{Lslope} = -0.43$ while the angry condition had a positive $\lambda_{Lslope} = 0.26$, but these two conditions did not show t-test significance.

154

5.3.2 Comparing different models for the PU data fit

Similarly to the model comparison for the EU data fit (Section 5.2.2), different models were compared to fit the self-reported PU data in terms of the BIC (Bayesian Information Criterion). It turned out that the M5 model (on which the dissertation focuses) was the best in terms of the BIC for PU just as it was for EU.

5.3.3 The domain, the PU frame and the emotion condition influence the probability of choosing the risky option

There were a main effect of the domain type ($M_{domain1} = 38\%$ vs. $M_{domain2} = 64\%$; $F(1,167) = 67.2$, $p < .001$), a main effect of the PU frame type ($M_{gain} = 47\%$ vs. $M_{loss} = 55\%$; $F(1,167) = 5.33$, $p < .05$) (supporting H7-1) and a marginal main effect of the emotion condition ($F(3,167) = 2.49$, $p = .06$) on the probability of choosing the risky option. Also, there was a significant interaction between the PU frame type and the emotion condition ($F(3,167) = 3.37$, $p < .05$).
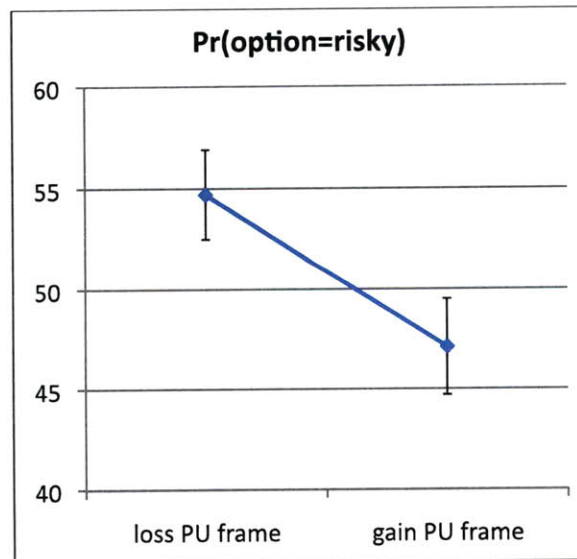
Figure 5-15: The main effect of the PU frame on the probability of selecting the risky option. Note that error bars indicate standard errors of the mean.

155
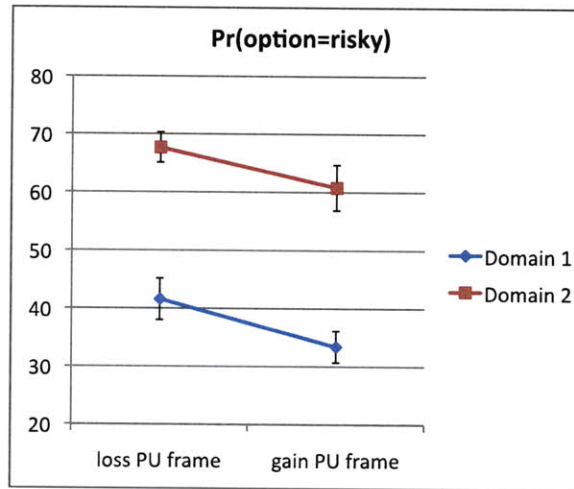
Figure 5-16: The interaction effect of the domain and the PU frame on the probability of selecting the risky option. Note that error bars indicate standard errors of the mean.
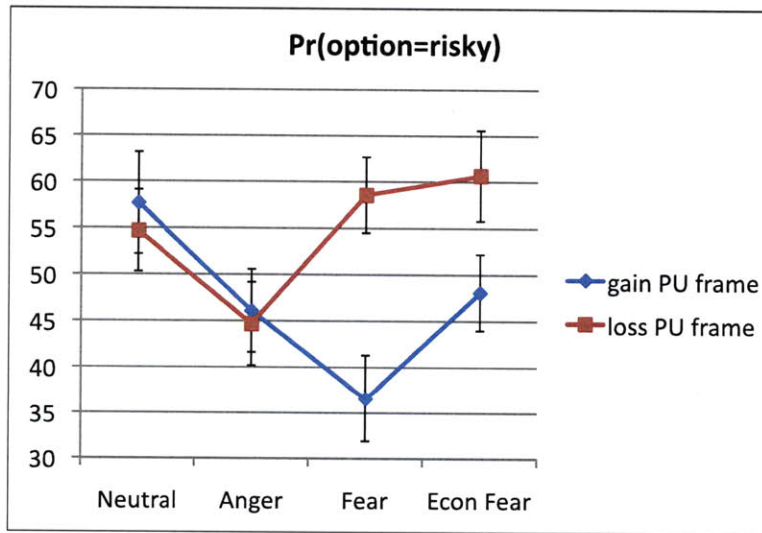
Figure 5-17: The interaction effect of the emotion and the PU frame on the probability of selecting the risky option. Note that error bars indicate standard errors of the mean.

The probability of choosing the risky option in the PU gain frame

When respondents employed the PU gain frame (i.e., viewing the expected outcomes of the safe option as gains), there were a main effect of the domain type ($M_{domain1} = 33\%$ vs. $M_{domain2} = 61\%$; $F(1,80) = 26.3$, $p < .001$), a marginal main effect of the emotion condition ($F(3,80) = 2.28$, $p = .08$) on the probability of choosing the risky option but

156

	Gain PU frame	Loss PU frame
Neutral	45% (N=18)	59% (N=24)
	$T(40) = -1.91, p < .05$	
Anger	41% (N=21)	50% (N=21)
	$T(40) = -1.21, NS$	
Fear	35% (N=18)	61% (N=24)
	$T(40) = -4.20, p < .001$	
Economic Fear	45% (N=24)	64% (N=18)
	$T(40) = -2.30, p < .05$	

Table 5.23: Comparing the probability of selecting the risky option for gain and PU loss frames on each emotion condition (pooling over two domains)

(a) Loss PU frame   (b) Gain PU frame
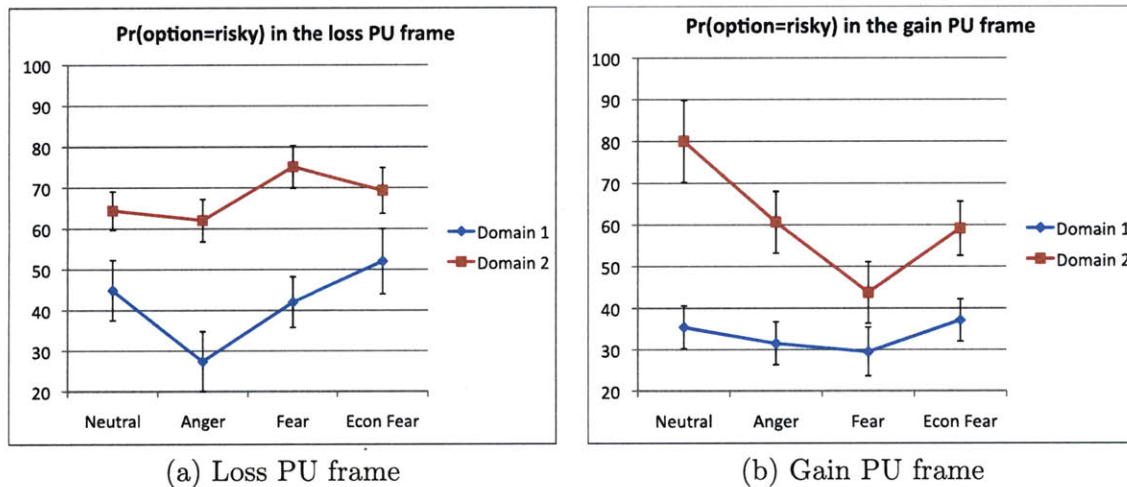
Figure 5-18: The interaction effect of the emotion and the domain on the probability of selecting the risky option for each PU frame (gain or loss). Note that error bars indicate standard errors of the mean.

no significant interaction between the domain type and the emotion condition ($F(3,80) = 1.24, NS$). Pairwise comparisons (LSD) of the estimated marginal means of each emotion condition showed that neural respondents selected the risky option significantly more often than fearful respondents in the PU gain frame ($p < .05$).

- Domain 1, PU gain frame: When respondents used the PU gain frame on the Domain 1 task, pairwise T-tests showed that there were no significant differences in the probability of selecting the risky option across emotional conditions.

- Domain 2, PU gain frame: Using the PU gain frame on the Domain 2 task, fearful respondents selected the risky option significantly less often than neutral respondents (M = 44% vs. 80%; $T(9)$ = -4.43, $p < .01$). Also, economically fearful respondents selected the risky option significantly less often than neutral respondents (M = 59% vs. 80%; $T(9.8)$ = -2.2, $p < .05$).

The probability of choosing the risky option in the PU loss frame

When respondents employed the PU loss frame (i.e., viewing the expected outcomes as losses), there were main effects of the domain type ($M_{domain1}$ = 42% vs. $M_{domain2}$ = 68%; $F(1,86)$ = 45.1, $p < .001$) and the emotion condition ($F(3,86)$ = 3.20, $p < .05$) on the probability of choosing the risky option but no significant interaction between them ($F(3,86)$ = 1.31, NS). Pairwise comparisons (LSD) of the estimated marginal means of each emotion condition showed that fearful and economically fearful respondents selected the risky option significantly more often than angry respondents in the PU loss frame ($ps < .01$). Also, neutral respondents selected the risky option marginally more often than angry respondents in the PU loss frame ($p = .07$).

- Domain 1, PU loss frame: For the PU loss frame on the Domain 1 task, angry respondents selected the risky option significantly less often than neutral respondents (M = 27% vs. 45%; $T(12)$ = -2.55, $p < .05$), fearful respondents (M = 27% vs. 42%; $T(15)$ = -1.8, $p < .05$) and economically fearful respondents (M = 27% vs. 52%; $T(11)$= -2.2, $p < .05$).

- Domain 2, PU loss frame: For the PU loss frame on the Domain 2 task, fearful respondents selected the risky option significantly more often than neutral respondents (M = 75% vs. 64%; $T(29)$ = 2.07, $p < .05$) and angry respondents (M = 75% vs. 62%; $T(26)$ = 2.29, $p < .05$).

5.3.4 The domain, the PU frame and the emotion condition influence the average PU difference of the risky option and the safe option (= Average[PU(risky option)] − Average[PU(safe option)])

Note that, by this definition, when the average PU difference is positive, the average PU of the risky option is greater than that of the safe option by the definition.

There were a main effect of the domain type ($M_{domain1}$ = -13.2 vs. $M_{domain2}$ = 8.5; $F(1,167) = 36.5$, $p <.001$) and a main effect of the PU frame type (M_{gain} = -10.6 vs. M_{loss} = 5.9; $F(1,167) = 20.9$, $p <.001$) on the average PU difference over all trials (PUrisk − PUsafe = Average[PU(risky option)] − Average[PU(safe option)]) (supporting H8-1). Also, there was a marginal interaction between the domain type and the PU frame type ($F(3,167)$ = 3.82, $p = .053$).
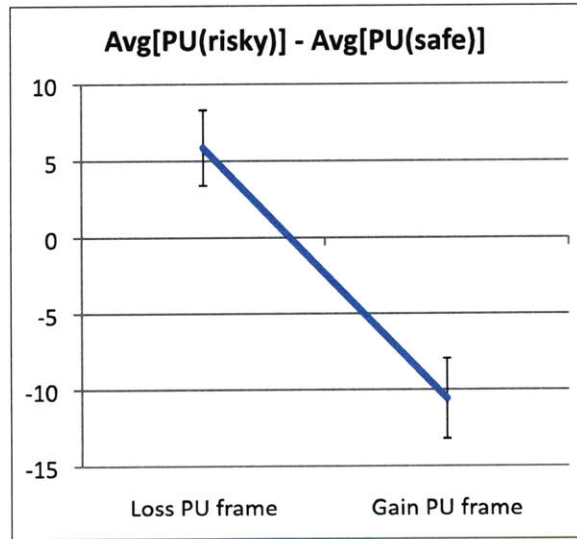
Figure 5-19: The main effect of the PU frame on the average PU difference. Note that error bars indicate standard errors of the mean.

The average PU difference (PUrisk − PUsafe) over all trials in the PU gain frame

When respondents employed the PU gain frame (i.e., viewing the expected outcomes of the safe option as gains), there were a main effect of the domain type ($M_{domain1}$ = -
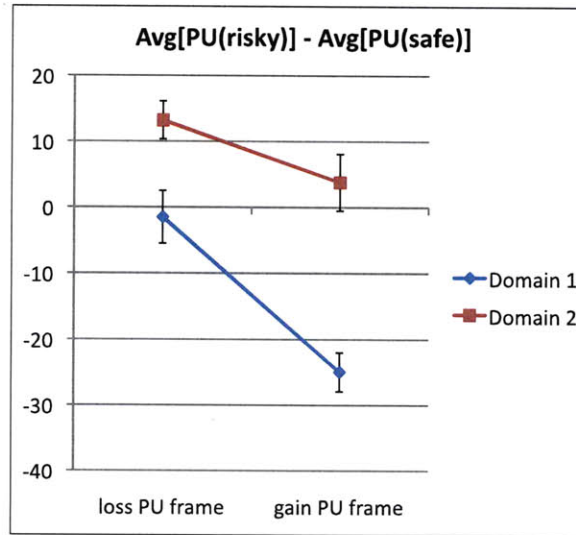
Figure 5-20: The interaction effect of the domain and the PU frame on the average PU difference. Note that error bars indicate standard errors of the mean.
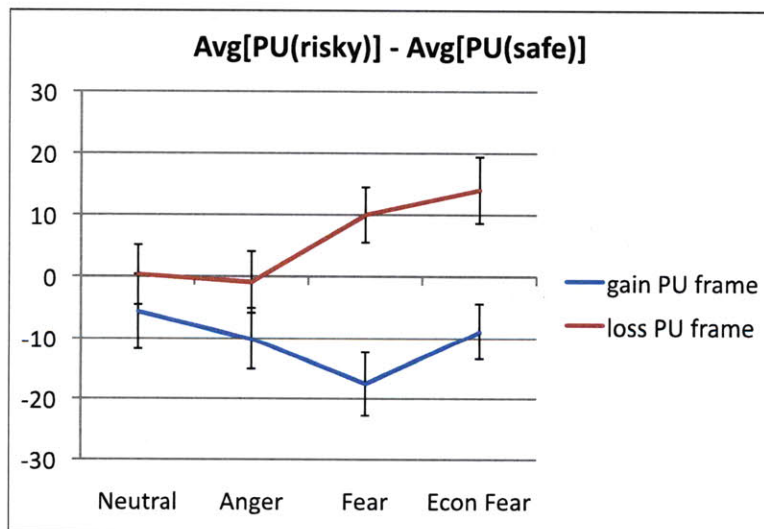
Figure 5-21: The interaction effect of the emotion and the PU frame on the average PU difference. Note that error bars indicate standard errors of the mean.

24.9 vs. $M_{domain2} = 3.8$; $F(1,80) = 24.8$, $p < .001$) on the average PU difference (= Average[PU(risky option)] − Average[PU(safe option)]), but no main effect of the emotion condition ($F(3,80) = .71$, NS) but no significant interaction between the domain type and the emotion condition ($F(3,80) = .41$, NS).

	Gain PU frame	Loss PU frame
Neutral	45% (N=18)	59% (N=24)
	$T(40) = -1.91, p < .05$	
Anger	41% (N=21)	50% (N=21)
	$T(40) = -1.21, NS$	
Fear	35% (N=18)	61% (N=24)
	$T(40) = -4.20, p < .001$	
Economic Fear	45% (N=24)	64% (N=18)
	$T(40) = -2.30, p < .05$	

Table 5.24: Comparing the average PU difference (= Average[PU(risky option)] − Average[PU(safe option)]) for gain and PU loss frames on each emotion condition (pooling over two domains)

(a) Loss PU frame   (b) Gain PU frame
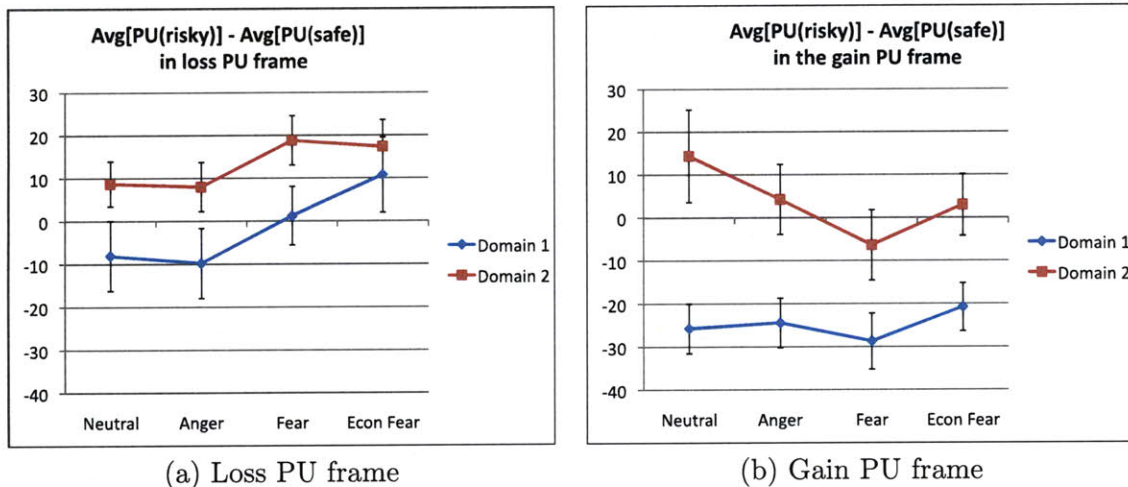
Figure 5-22: The interaction effect of the emotion and the domain on the average PU difference for each PU frame (gain or loss). Note that error bars indicate standard errors of the mean.

- Domain 1, PU gain frame: When respondents used the PU gain frame on the Domain 1 task, pairwise T-tests showed that there were no significant differences across emotional conditions in terms of the average PU difference of two options.

- Domain 2, PU gain frame: Using the PU gain frame on the Domain 2 task, fearful respondents reported negative average PU difference significantly smaller than neutral respondents ($M = -6.4$ vs. 14.4; $T(9) = -2.55, p < .05$).

161

The average PU difference (PUrisk – PUsafe) over all trials in the PU loss frame

When respondents employed the PU loss frame (i.e., viewing the expected outcomes of the safe option as losses), there were a main effect of the domain type ($M_{domain1}$ = -1.5 vs. $M_{domain2}$ = 13.2; $F(1,86) = 11.0, p < .001$) and a marginal main effect of emotion condition ($F(3,86) = 2.63, p = .06$) on the average PU difference (= Average[PU(risky option)] – Average[PU(safe option)]) but no significant interaction between them ($F(3,86) = .33, NS$). Pairwise comparisons (LSD) of the estimated marginal means of each emotion condition showed that economically fearful respondents reported significantly greater average PU difference than neutral and angry respondents in the PU loss frame ($ps < .01$). Also, fearful respondents reported marginally greater average PU difference than angry respondents ($p = .07$) and neutral respondents ($p = .08$) in the PU loss frame.

- Domain 1, PU loss frame: For the PU loss frame on the Domain 1 task, angry respondents reported marginally smaller negative average PU difference (i.e., more risk averse attitude) than fearful respondents and economically fearful respondents ($ps = .06$).

- Domain 2, PU loss frame: For the PU loss frame on the Domain 1 task, fearful respondents reported marginally greater average PU difference (i.e., more risk seeking attitude) than neutral respondents ($M = 18.8$ vs. 8.7; $T(29)=1.50, p = .07$).

5.4 Manipulation checks

At the end of the experiment, respondents were asked to report how they felt while writing a diary and watching each video clip (video 1, video 2, video 3). To avoid revealing our interest in specific emotions, we included 8 affective states ("angry", "anxious", "disgusted", "fearful", "happy", "interested", "irritated", "sad") on the form, although only 4 were of interest. An anger factor included "anger" and "irritated" ($\alpha = .903$), and a fear factor included "anxious" and "fearful" ($\alpha = .942$). Response scales ranged from 0 (do not feel the emotion the slightest bit) to 8 (feel the emotion even more strongly than ever before).

162

We averaged responses on each emotion for subsequent analyses.

There were four manipulation steps (diary, video 1, video 2, video 3). To control for individual biases, each individual's self-reported rating on an affective state was standardized using his or her average rating and standard deviation across all affective states over all manipulation steps:

For the diary step, individual analysis of variance (ANOVAs) on self-reported experience of anger, $F(3,80) = 15.5$, and fear $F(3,80) = 12.1$, revealed strong emotion-induction effects ($ps < .001$). Post-hoc comparisons (LSD) for anger showed that the anger condition ($M = .598$) induced significantly more anger than the neutral condition ($M = -.379$) and the fear condition ($M = -.385$) ($ps < .001$). Post-hoc comparisons (LSD) for fear showed that the fear condition ($M = .369$) induced significantly more fear than the neutral condition ($M = -.207$) and the anger condition ($M = -.355$) ($ps < .05$). The economic fear condition induced significantly more anger and fear than the neutral condition ($ps < .05$). In addition, respondents felt significantly more anger than fear in the anger condition ($M = .598$ vs. $-.355$; $T(20) = 5.228$, $p < .001$), and significantly more fear than anger in the fear condition ($M = .369$ vs. $-.385$; $T(20) = 5.125$, $p < .001$). Respondents didn't feel significantly more fear than anger in the economic fear condition, although the mean level of fear was greater than that of anger.

Figure 5-23: manipulation check: diary

For the video 1 step, individual analysis of variance (ANOVAs) on self-reported experi-

ence of anger, $F(3,80) = 11.3$, and fear $F(3,80) = 15.7$, revealed strong emotion-induction effects $(ps < .001)$. Post-hoc analysis for anger showed that the anger condition $(M = .551)$ induced significantly more anger than the neutral condition $(M = -.411)$ and the fear condition $(M = -.255)$ $(ps < .001)$. Post-hoc analysis for fear showed that the fear condition $(M = .787)$ induced significantly more fear than the neutral condition $(M = -.368)$ and the anger condition $(M = -.138)$ $(ps < .001)$. The economic fear condition induced significantly more anger $(ps < .05)$ and fear $(ps < .001)$ than the neutral condition. In addition, respondents felt significantly more anger than fear in the anger condition $(M = .551$ vs. $-.138$; $T(20) = 4.423$, $p < .001)$, and significantly more fear than anger in the fear condition $(M = .787$ vs. $-.255$; $T(20) = 4.688$, $p < .001)$. Respondents didn't feel significantly more fear than anger in the economic condition, although the mean level of fear was greater than that of anger.

Figure 5-24: manipulation check: video1

For the diary and video 1 steps overall, individual analysis of variance (ANOVAs) on self-reported experience of anger, $F(3,80) = 21.7$, and fear $F(3,80) = 21.2$, revealed strong emotion-induction effects $(ps < .001)$. Post-hoc analysis for anger showed that the anger condition $(M = .575)$ induced significantly more anger than the neutral condition $(M = -.395)$ and the fear condition $(M = -.320)$ $(ps < .001)$. Post-hoc analysis for fear showed that the fear condition $(M = .578)$ induced significantly more fear than the neutral condition $(M = -.287)$ and the anger condition $(M = -.247)$ $(ps < .001)$. The economic fear condition

164

induced significantly more anger ($ps < .005$) and fear ($ps < .001$) than the neutral condition. In addition, respondents felt significantly more anger than fear in the anger condition ($M = .575$ vs. -.247; $T(20) = 6.773$, $p < .001$), and significantly more fear than anger in the fear condition ($M = .578$ vs. -.320; $T(20) = 5.876$, $p < .001$). Respondents didn't feel significantly more fear than anger in the economic condition, although the mean level of fear was greater than that of anger.
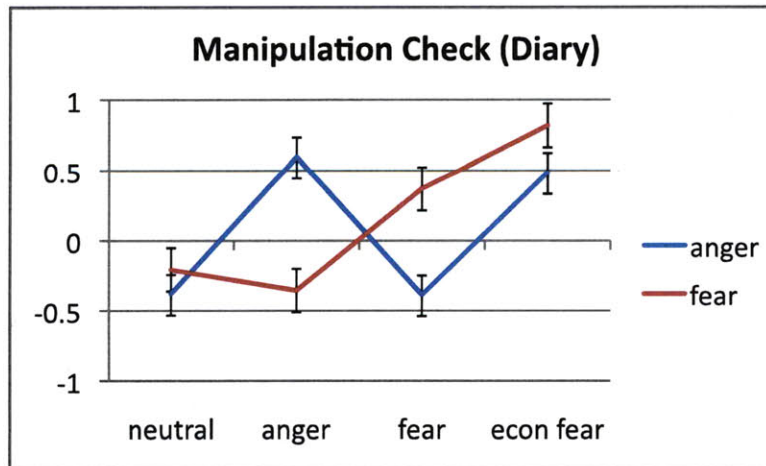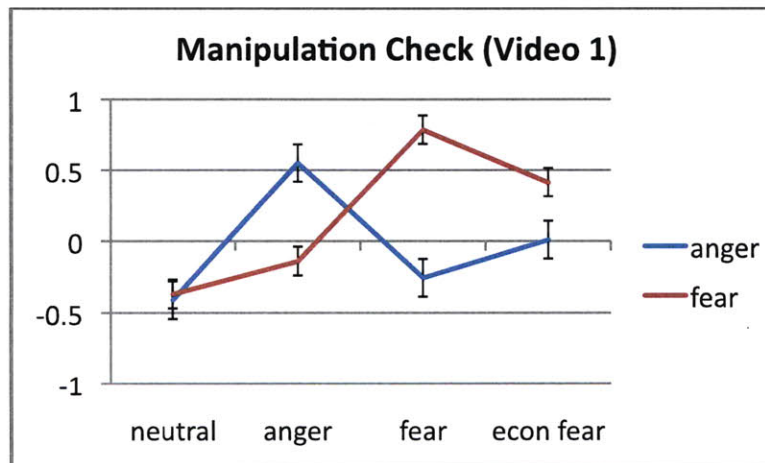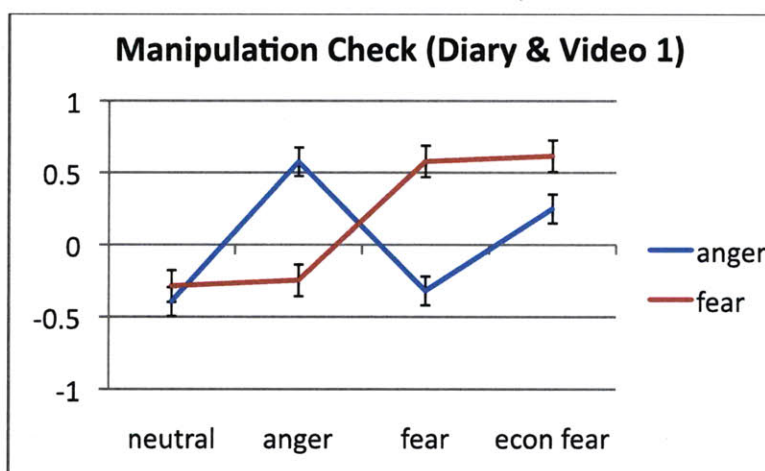

Figure 5-25: manipulation check: diary and video1

For the video 2 step, individual analysis of variance (ANOVAs) on self-reported experience of anger, $F(3,80) = 22.7$ ($p < .001$), and fear $F(3,80) = 4.80$ ($p < .005$), revealed strong emotion-induction effects. Post-hoc analysis for anger showed that the anger condition ($M = .818$) induced significantly more anger than the neutral condition ($M = -.382$) and the fear condition ($M = -.389$) ($ps < .001$). Post-hoc analysis for fear showed that the fear condition ($M = .487$) induced significantly more fear than the neutral condition ($M = -.231$, $p < .001$). The economic condition induced significantly more anger ($p < .05$) and fear ($p < .005$) than the neutral condition. In addition, respondents felt significantly more anger than fear in the anger condition ($M = .818$ vs. .262; $T(20) = 2.866$, $p < .01$), and significantly more fear than anger in the fear condition ($M = .487$ vs. -.389; $T(20) = 5.207$, $p < .001$). Respondents felt more fear than anger in the economic condition ($M = .442$ vs. .026; $T(20) = 1.99$, p = .06).

For the video 3 step, individual analysis of variance (ANOVAs) on self-reported expe-
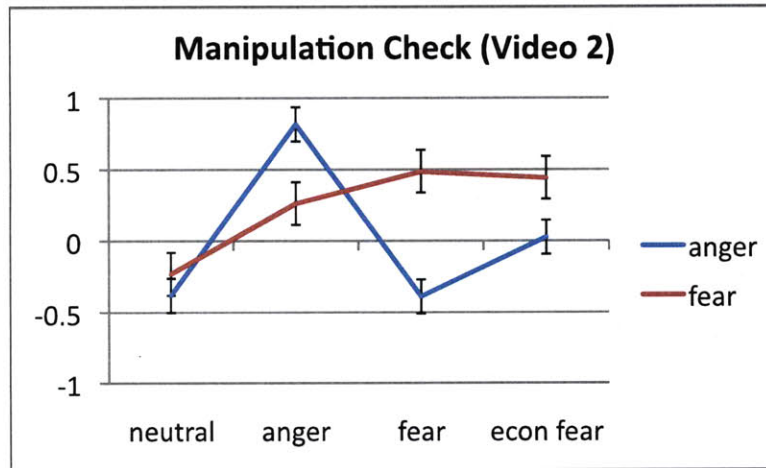
Figure 5-26: manipulation check: video2

rience of anger, $F(3,80) = 7.390$ $(p < .001)$, and fear $F(3,80) = 8.056$ $(p < .001)$, revealed strong emotion-induction effects. Post-hoc analysis for anger showed that the anger condition $(M = .474)$ induced significantly more anger than the neutral condition $(M = -.357)$ and the fear condition $(M = -.410)$ $(ps < .001)$. Post-hoc analysis for fear showed that the fear condition $(M = .051)$ induced significantly more fear than the neutral condition $(M = -.429)$ and the anger condition $(M = -.326)$ $(ps < .05)$. The economic condition induced significantly more fear $(p < .001)$ than the neutral condition. In addition, respondents felt significantly more anger than fear in the anger condition $(M = .474$ vs. $-.326$; $T(20) = 4.780$, $p < .001)$, and significantly more fear than anger in the fear condition $(M = .051$ vs. $-.410$; $T(20) = 2.286$, $p < .05)$. Respondents felt more fear than anger in the economic condition $(M = .377$ vs. $-.094$; $T(20) = 2.03$, p $= .06)$.

Overall, it appears that the protocol of repeatedly eliciting a state of fear or of anger worked well for the pure anger and pure fear states, for all videos and diary elicitors. The economic fear condition tended to elicit both fear and anger, although more fear than anger.

166
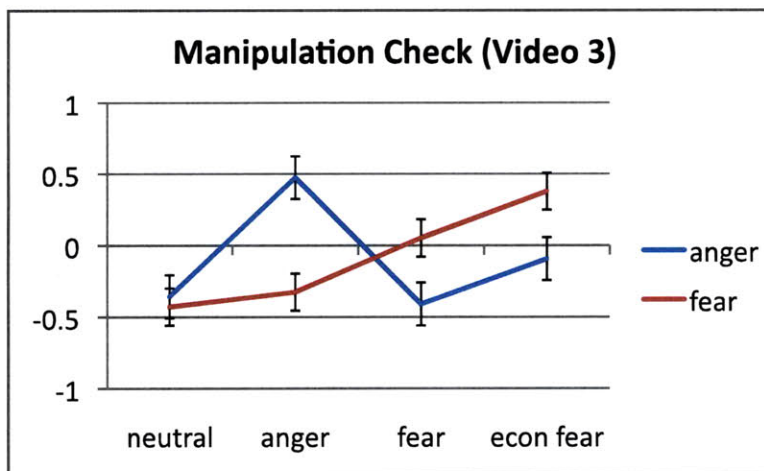
Figure 5-27: manipulation check: video3

Chapter 6

Conclusion and Discussion

This dissertation presented a new computational perspective on the role of subjective experience and affective prediction in human decision making and learning, drawing on the findings in diverse areas of decision science as behavioral economics, neuroeconomics, psychology and machine learning.

6.1 Summary

In summary, the dissertation

- Defined subjective discriminability. Showed both computationally and empirically that bigger subjective discriminability leads to more optimal decisions.

- Characterized how subjective and affective influences may help or harm human decision making depending on domains, frames, emotions and their interactions.

- Constructed a new model combining measures to evaluate risk preferences: behavioral choices, self-reported experience (subjective discriminability), self-reported predicted utility (predicted-utility difference), self-reported confidence.

- Introduced two different kinds of subjective value functions (experienced-utility (EU) function and predicted-utility (PU) function) whose parameters change with emotions and provided a method to infer PU and EU functions from self-reported EU and PU data in each emotion condition.

- Showed how to compute reference points (EU frame, PU frame) for each of the EU and PU functions.

- Analyzed risk attitudes based on EU and PU frames as well as on the frame given by experimenter.

- Observed how emotions influence the reference point selection (framing).

- Discovered the frame and emotion effects (main and interaction effects) in decision making under uncertainty.

- Measured how experience, gender and emotion influence confidence and prediction.

- Introduced the confidence-dependent predicted utility function.

- Presented a new emotion-refresher method.

- Defined the value of risk (VOR).

- Characterized how domain, frame, emotion influence decision making: Negative emotions in face of gains (more risk-averse), Anger in face of losses (more risk-averse), Fear and Economic fear in face of losses (more risk-seeking).

- Showed how human behavior can be described by emotion-shaped EU and PU functions.

- Provided a theory that better explain/simulate human behavior under uncertainty, frames and emotions.

6.2   Simulations and Empirical Results

Subjective and affective elements are well-known to influence human learning and decision making. The research for exploring and exploiting these important influences in computational learning theory, however, is still in its early stage. This section presents a new model combining subjective and affective influences within the RL and MDP framework. The affective-congitive (AC) model involves two different modes: the experience-based mode and

the prediction-based mode. To model the total-experienced utility from past experiences, the AC model introduces a prospect theory (PT)-based parameterized "experienced-utility function". In order to model affective-subjective characteristics of prediction-based mode, the AC model employs a prospect theory (PT)-based parameterized "predicted-utility function". It also models one specific kind of affective state, called the "goal-achieving (confidence) state," which relates to the sense of confidence in the current decision-making policy. In economics theory, the PT value function is fixed, but it is hypothesized that the affective state influences the shape of the predicted-utility function (i.e., sensitivities to the expected gains and losses). An RL-based computational framework that implements this hypothesis automatically regulates trade-offs between exploration and exploitation while beating the performance of five other well-known model-free learning algorithms.

The AC model includes both a subjective component (PT value function) and a component that captures part of how affective states may influence decision making. It is further hypothesized that the latter component can influence the former to give performance that is closer to human behavior. Furthermore, while PT theory has been developed in economics for domains with known outcome distributions, this new model enables PT theory to be used for unknown and changing stochastic outcome distributions. Finally, it is known that in the face of multiple unknown nonstationary distributions of outcomes, balancing the trade-offs between exploration and exploitation is very critical; this new model achieves this balanced trade-off in an automatic and internally-regulated way.

Human decision-making experiments were conducted to empirically infer how people adjust the parameters (i.e., risk attitude and reference point) of their experienced-utility and predicted-utility functions in sequential decision-making situations involving incidental affective states (e.g., anger, fear) and task-related affects (e.g., confidence). Computational simulations confirmed that the location of the reference point was very critical in optimal decision making. The same framing effects were observed in human behavioral experiments.

Barberis et al. [2] showed how previous outcomes could influence the risky choice in terms of the slope change of a piecewise-linear approximation to the traditional PT function. They assume that previous gains decrease the sensitivity to potential losses and previous losses increase the sensitivity to potential losses. This is different from the assumption of

the confidence-based predicted-utility (PU) function in the AC model. The AC model hypothesized that previous gains lead to higher confidence, which increases the PU function's sensitivities to both estimated gains and losses. The higher sensitivity then increases the discriminability (PU difference) between two options, facilitating the exploitative choice in the next trial. Previous losses, however, lead to lower confidence, which decreases the PU function's sensitivities to both estimated gains and losses. The lower sensitivity then decreases the discriminability (PU difference) between two options, facilitating the exploratory choice in the next trial. The confidence-based hypothesis helped the AC model control the trade-offs between exploration and exploitation in simulations. The empirical evaluations from human experiments showed that previous gains (losses) led to higher (lower) confidence, significantly increasing (decreasing) the sensitivity to estimated gains, respectively. There was, however, no significant changes on the sensitivity to estimated losses.

People evaluate their outcome relative to their reference point. Although I assumed that the reference point is fixed over trials during a task for the analysis, the reference point may change over trials during a task. It will be very interesting to investigate how people change their reference point over trials in a stationary or an unpredictable dynamic multi-armed bandit tasks. Also, it will be worthwhile to infer the optimal number of trials for initial exploration and see how the number is related to their selection of the reference point.

The affective sensor data (facial valence and skin conductance arousal) measured during the experiment will be analyzed in the future and might be used as an objective measure of the experienced utility. It will be interesting to see if there are any differences between the self-reported experienced-utility and the sensor-measured experienced-utility (e.g., in a bandit task involving repetitive trials, the sensor-measured implicit utility might be an earlier indicator of their liking/disliking of a certain choice than the self-reported explicit utility).

Regarding the emotion manipulation in human experiments, the same video clips were used in two different steps of emotion manipulation (the first and the third emotion-manipulation steps in neutral, anger and fear conditions) to refresh people's emotion condition. The manipulation check through the after-the-test questionnaire confirmed that the new emotion-refresher method successfully induced similar levels of emotion responses and

refreshed the emotion condition.

Figure 5-10 presents the proposed model on the interactions of gender, emotion, experience, confidence, prediction and exploitative/exploratory decisions. Gender and emotion (e.g., incidental fear) may influence confidence. Good (bad) experience leads to high (low) confidence. Also, then, high (low) confidence leads to high (low) predicted-utility difference and more exploitative (exploratory) choice behavior. That is, when respondents have a relatively good experience, they get more confident in the task, compared to when they have a relatively bad experience. When respondents are more confident in the task, they tend to predict that one option is much better than the other, compared to when they are less confident. When respondents tend to predict that one option is better than the other, they are more likely to select the better option (= exploitative choice).

Interestingly, female respondents in the fear condition reported significantly lower average confidence than male respondents in that condition. Fearful female respondents tended to feel less confident even with previous (relatively) higher experienced utility, possibly due to the influence of the existing incidental fearful emotion. Moreover, since lower confidence led to lower predicted-utility (PU) difference and lower PU difference led to more randomized choice behavior, fearful female respondents tended to have more exploratory trials (i.e, more randomized choices) than fearful male respondents. In other words, fearful female respondents selected fewer exploitative choices (i.e., fewer selections of the option with the current-estimated greatest predicted utility) than fearful male respondents.

Consider a two-armed bandit task (such as Domain 1 or Domain 2) in which each option involves a different outcome variance. For reinforcement-learning algorithms with an exploration strategy such as the value of information (VOI), the option with a greater outcome variance is chosen (or explored) more often than the one with a smaller uncertainty. This sort of exploration strategy might be linked to curiosity-dependent human behavior (e.g., curiosity for an option may be modeled by the entropy of the posterior outcome distribution).

Also, some well-known economic models of choice such as the Markowitz-Tobin portfolio selection model make a trade-off between mean (μ_x) and outcome variance (σ_x^2) in computing the expected utility of an option [48]: expected utility $(X) = \mu_x - a\sigma_x^2$ where X denotes a

random outcome variable of the option and a (> 0) is the risk-aversion coefficient. Thus, as the outcome variance of an option becomes greater, the choice preference for that option becomes lower.

In contrast, I hypothesize in the AC model that human decision behaviors are influenced not only by the attitude to maximize long-run average outcome, but also by the risk attitude associated with the selected reference point.

- Under the gain frame, people would not prefer the option with a high outcome variance to its certainty equivalent. As the outcome variance of an option becomes greater, people tend to show less preference for that option (risk aversion in the gain frame). People are risk averse when it comes to gains.

- Under the loss frame, people would prefer the option with a high outcome variance to its certainty equivalent. As the outcome variance becomes greater, people tend to show more preference for that option (risk seeking in the loss frame). People are more willing to gamble when it comes to losses.

Compared with the value of information (VOI), I will conceptualize the value of risk (VOR) regarding the AC model. VOR measures the risk attitude for a stochastic outcome distribution. The VOR for an outcome distribution can be defined as the difference between the subjective value of the outcome distribution (X) and that of its certainty-equivalent (μ_x): $\mathrm{VOR} = f(X) - f(\mu_x)$ where the subjective value function f can be either the predicted-utility (PU) function or the experienced-utility (EU) function. Note that the value of $f(X)$ depends not only on μ_x and σ_x^2 but also on the reference point selection, and that VOR < 0 (risk aversion) in the gain frame and VOR > 0 (risk seeking) in the loss frame. That is, the VOR is not absolute, but relative to the location of the reference point. The analysis results of human experiments on peoples risk attitudes supported the hypothesis that the VOR critically influences respondents' choice behavior. That is, the VOR explains framing effects.

Even for the same outcome distribution, people would have different levels of risk attitudes according to their information on outcome distributions or their availability of sampling experiences. Shafir et al. [52] have suggested that perceptual accuracy and dis-

174

criminability in experience-based decisions influence risk-taking behavior. That is, under the gain frame, the risk-averse tendency (certainty effect, people's tendency to select the safer of two prospects) emerges when it is difficult to discriminate between the different outcomes (e.g., hard-to-assess outcomes in the form of graphical dots), whereas the risk-averse tendency disappears when the discrimination is easy (e.g., easy-to-assess outcomes in the form of digital numbers). Althoug they focused on the perceptual discriminability involving different forms of outcomes, it would be very interesting to examine how the subjective discriminability as defined in this dissertation influences peoples risk attitudes in their experience-based decision behavior. For instance, in experience-based decisions, people might show evident risk attitudes for low-discriminability cases, but might not show risk attitudes for high-discriminability cases. In other words, in high-discriminability cases, it might be too easy for people to judge which option is better than the other, thus the choices might not be critically influenced by risk attitudes.

The incidental emotion state can influence the risk attitude or the parameter values of the PT-based experienced-utility and predicted-utility functions. I conducted human two-armed bandit experiments to infer peoples predicted utility and experienced utility functions under different incidental emotion states. My hypotheses on these functions with fear and anger in Section 2.5.7 were based on Lerner and Keltner's studies [32, 34, 35]. I hypothesized that the risk attitudes found in PT description-based predictions would appear in experience-based decisions. In my experiment results, however, there were significant interactions between framing and emotion. In the loss frame, anger made people more risk averse and fear made people more risk seeking. This seems to go against a wide range of findings.

At first this seems to go agoanst the existing findings; however, my conditions are actually different in an important way. For example, Lerner and Keltner's paper [34] focuses on how dispositional anger and fear influence people's prediction-based risk attitude (one-shot prediction such as in the prospect theory experiments) but my dissertation focuses on how incidental anger and fear influence sequential adaptive decisions under uncertainty in which the role of experience is very critical [21, 44, 23]. The experience-based learning situations are more uncertain and arousing than the one-shot prediction situation, due to the lack of

information (we only told them one option has a higher mean than the other). Thus, this could make my results differ from others who examined the influences of anger and fear without arousal and without sequential experiences.

The experienced-utility (EU) analysis on respondents' self-reported EU data in Figure 5-7 (d) confirms that

- Angry respondents tended to be less sensitive to relatively small experienced losses (e.g., moderate losses from the safe option) and very sensitive to relatively large experienced losses (e.g., very big losses from the risky option), compared to fearful respondents. This tendency made angry respondents more risk averse after several loss experiences.

- Fearful respondents tended to be more sensitive to relatively small experienced losses, compared to angry respondents. Also, due to this greater sensitivity to relatively small experienced losses, fearful respondents may have felt very bad for moderate losses from the safe option. This eventually contributed to a more risk-seeking attitude of fearful respondents after several loss experiences.

One possible explanation for the anger shaping we are seeing may be that anger is often present with some amount of pain - usually psychological or physical (e.g. somebody hits you or irritates you repeatedly and your quite normal reaction is anger). Here the "some amount of pain" includes a small financial loss, and so you are not that sensitive to it while angry. But as you receive larger losses, you then begin to react much more strongly than you would in a neutral state - (enhanced sensitivity to losses). We see this shape in the responses we measured in all the cases where there was anger.

The same tendency was found more clearly from the self-reported predicted-utility (PU) data analysis. The analysis of self-reported predicted-utility (PU) data in Figure 5-13 (d) confirms that

- The PU function shape in the angry condition (greater b parameter and smaller λ_{Lbase} parameter) implies that angry respondents predicted lower displeasure for relatively small losses (e.g., outcomes from the safe option) and greater displeasure for relatively

greater losses (e.g., very bad losses from the risky option), compared to neutral and fearful respondents.

- The PU function shape in the fear condition (smaller b parameter and greater λ_{Lbase} parameter) implies that fearful respondents were very sensitive to potential small losses (e.g., outcomes from the safe option). This leads to more risk seeking in face of potential losses.

These findings from the self-reported PU data analysis are consistent with those from the self-reported EU data analysis. Angry people tended to be less sensitive to potential small losses (e.g., moderate losses from the safe option) and very sensitive to potential big losses (e.g., big losses from the risky option), whereas fearful people tended to be very sensitive to even potential small losses. These prediction tendencies in the face of likely losses contributed to a more risk-averse attitude of angry respondents (selecting the safe option more often) and a more risk-seeking attitude of fearful respondents (selecting the risky option more often).

Note that the findings from the cognitive measures (self-reported EU measure in Section 5.2.3 and PU measure in Section 5.3.4) were consistent with those from actually observed behavioral choice measure (i.e., probability of choosing the safe vs. risky options during the tasks) in Section 5.3.3.

The AC model can be effectively applied to a non-human system that involves the affect-like state and is affected by that state. For instance, the stock market has market sentiment, "the intuitive feeling of the investment community regarding the expected movement of the stock market" (Wikipedia, 2008). A bullish or bearish market sentiment would be indicated by expected rising or falling prices, respectively. When a stock market sentiment is bullish or bearish, investors sensitive to the market sentiment would be more risk-seeking or more risk-averse, respectively. These decision behaviors could be described by the change of investors predicted-utility function influenced by the market sentiment. Also, investors should have an appropriate level of sensitivity to market sentiment for their best investment decisions. Another interesting phenomenon in the stock market is associated with "disposition effect". This refers to the tendency of investors to sell winning stocks, while holding onto losing

177

stocks. When the stock price is increasing, investors are often eager to realize gains. Also, they often become too anxious (economically fearful) and predict that the stock price would drop soon. That is, they tend to view the current investment in terms of gain framing and be too risk-averse. However, when the stock price is decreasing, investors are often reluctant to realize losses. They are likely to view the current investment in terms of loss framing and more anxiety in loss framing will increase the risk-seeking tendency.

Appendix A

Experienced-utility (EU) function in each emotion condition

Figure A-1: Experienced-utility (EU) function in the neutral condition (pooling over respondents). The points show self-reported EUs: red circle (safe option) and blue circle (risky option) on Domain 1, red diamond (safe option) and blue diamond (risky option) on Domain 2.

Figure A-2: Experienced-utility (EU) function with the *inferred* EU reference point (shown by the magenta line) for each respondent in the neutral condition on Domain 1. The points show self-reported EUs: red circle (safe option) and blue circle (risky option). Note that the EU reference point is varied over respondents on Domain 1.

Figure A-3: Experienced-utility (EU) function with the *inferred* EU reference point (shown by the magenta line) for each respondent in the neutral condition on Domain 2. The points show self-reported EUs: red diamond (safe option) and blue diamond (risky option). Note that the EU reference point is varied over respondents on Domain 2.
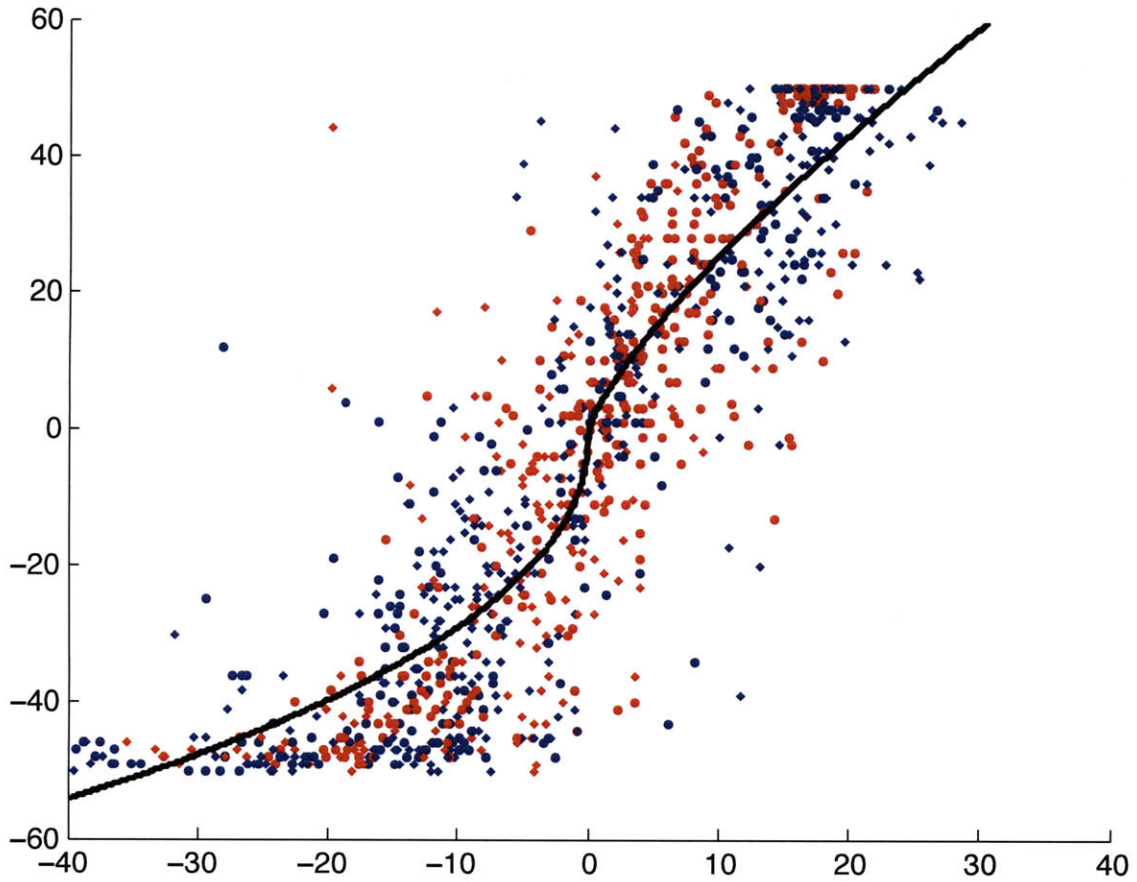
Figure A-4: Experienced-utility (EU) function in the anger condition (pooling over respondents). The points show self-reported EUs: red circle (safe option) and blue circle (risky option) on Domain 1, red diamond (safe option) and blue diamond (risky option) on Domain 2.
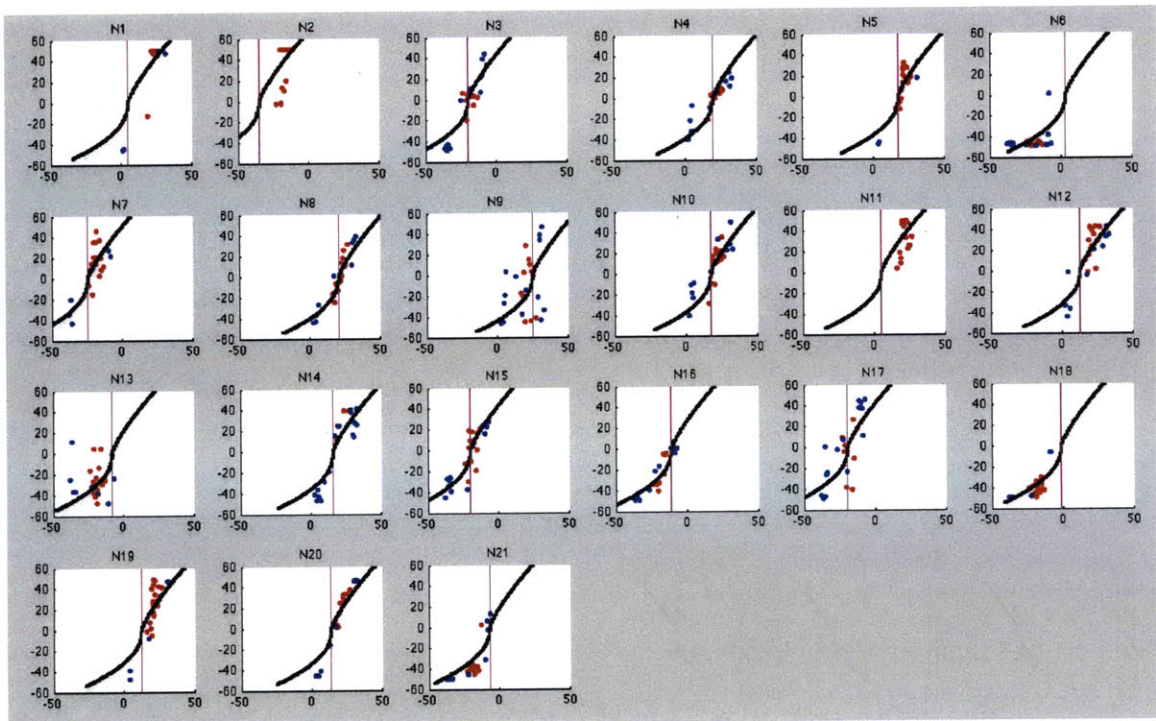
Figure A-5: Experienced-utility (EU) function with the *inferred* EU reference point (shown by the magenta line) for each respondent in the anger condition on Domain 1. The points show self-reported EUs: red circle (safe option) and blue circle (risky option). Note that the EU reference point is varied over respondents on Domain 1.
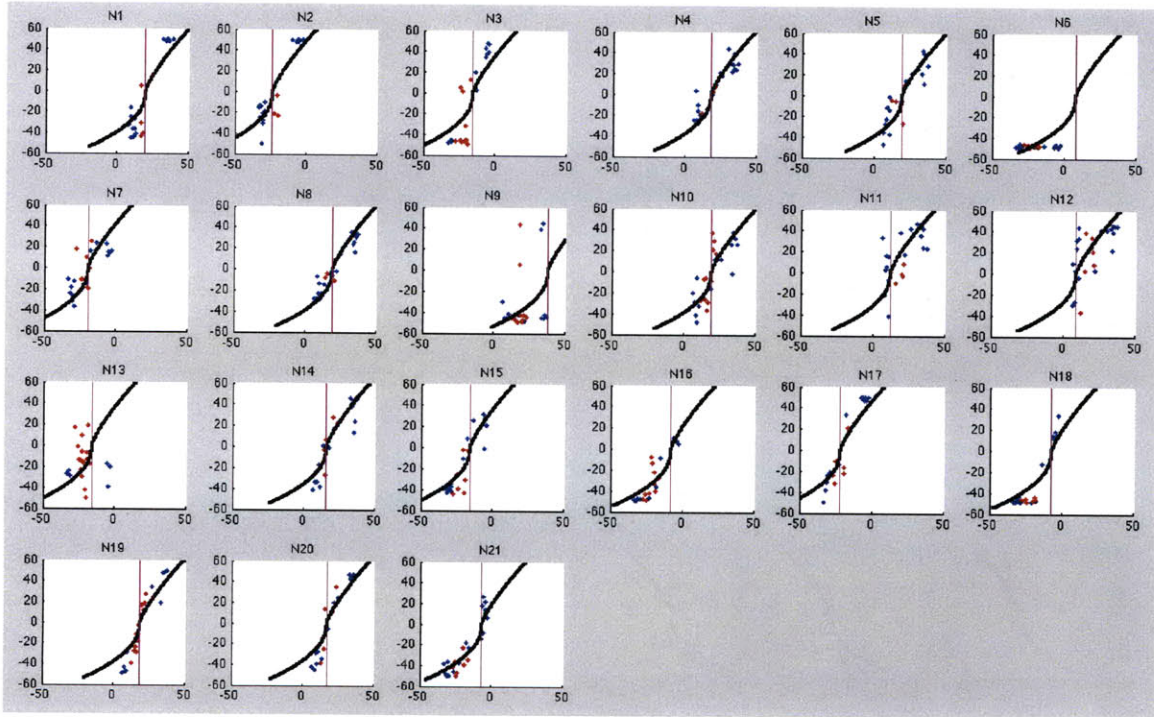
Figure A-6: Experienced-utility (EU) function with the *inferred* EU reference point (shown by the magenta line) for each respondent in the anger condition on Domain 2. The points show self-reported EUs: red diamond (safe option) and blue diamond (risky option). Note that the EU reference point is varied over respondents on Domain 2.
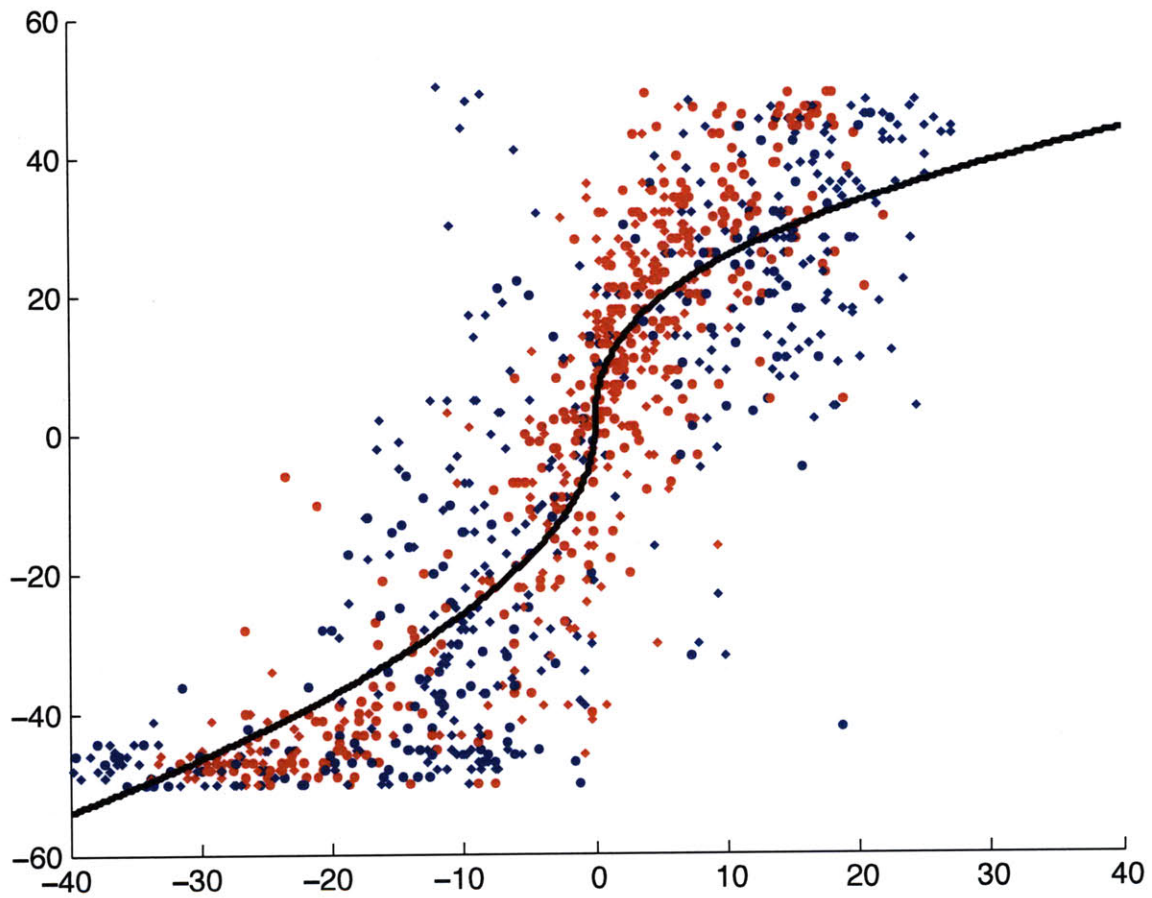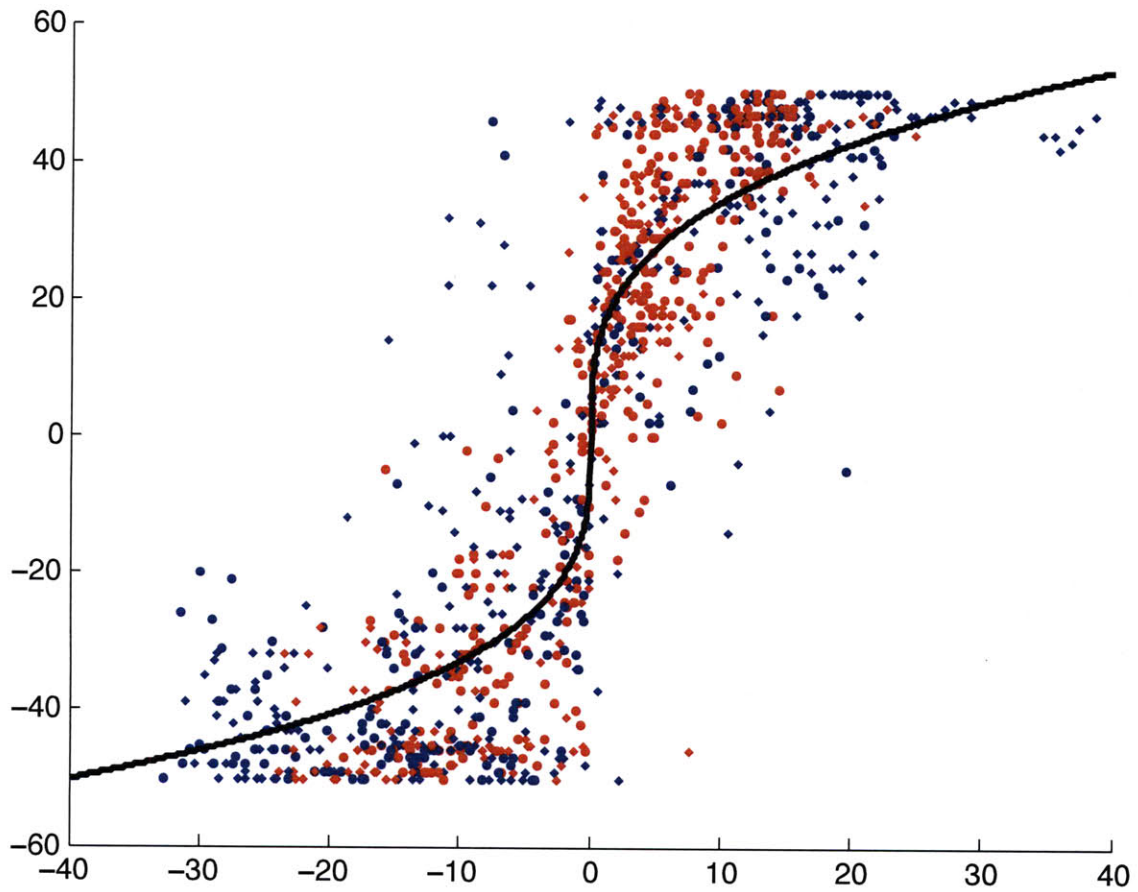
Figure A-7: Experienced-utility (EU) function in the fear condition (pooling over respondents). The points show self-reported EUs: red circle (safe option) and blue circle (risky option) on Domain 1, red diamond (safe option) and blue diamond (risky option) on Domain 2.
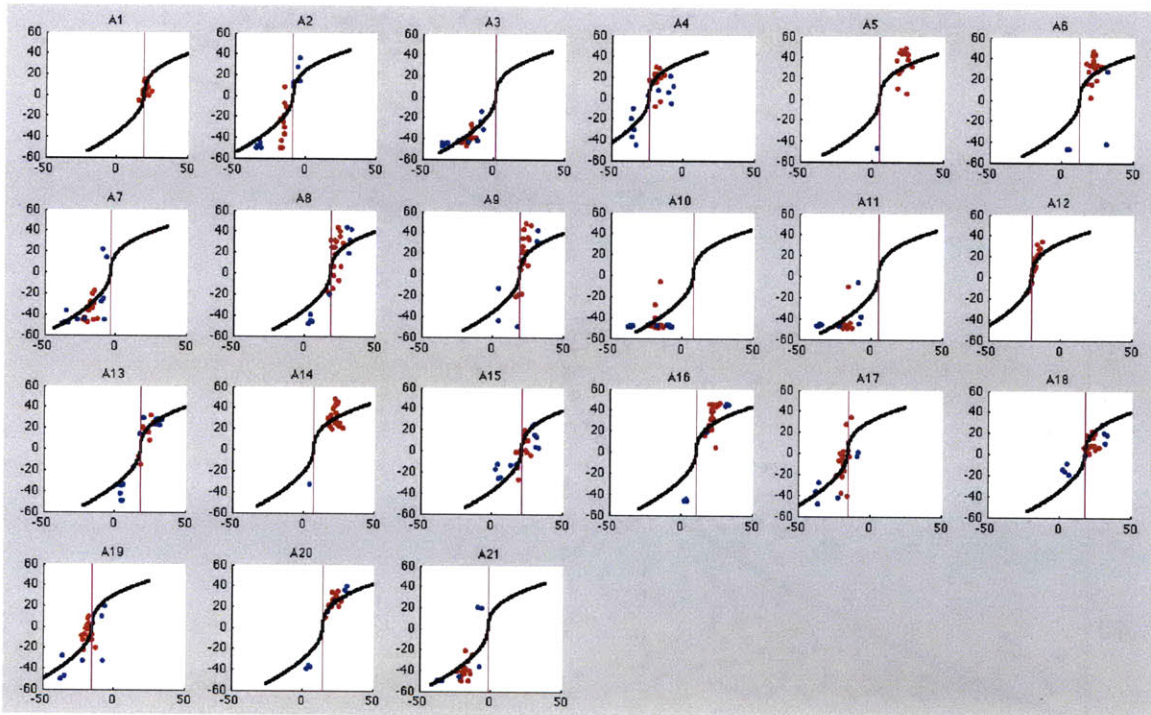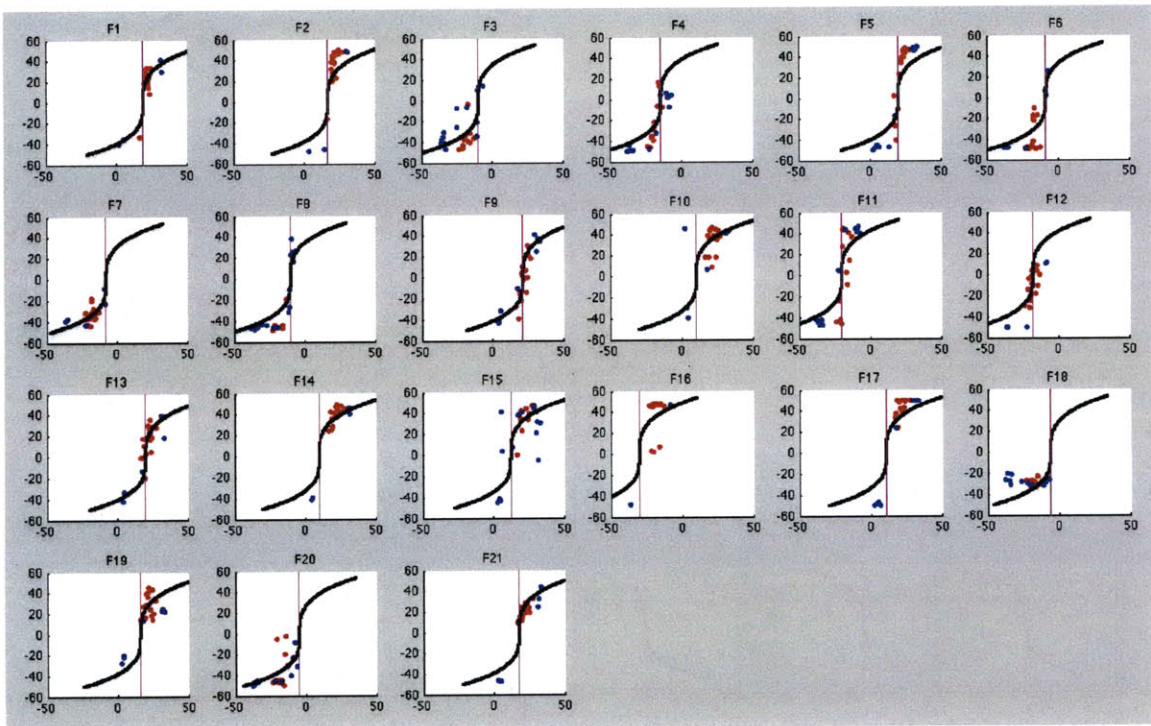
Figure A-8: Experienced-utility (EU) function with the *inferred* EU reference point (shown by the magenta line) for each respondent in the fear condition on Domain 1. The points show self-reported EUs: red circle (safe option) and blue circle (risky option). Note that the EU reference point is varied over respondents on Domain 1.
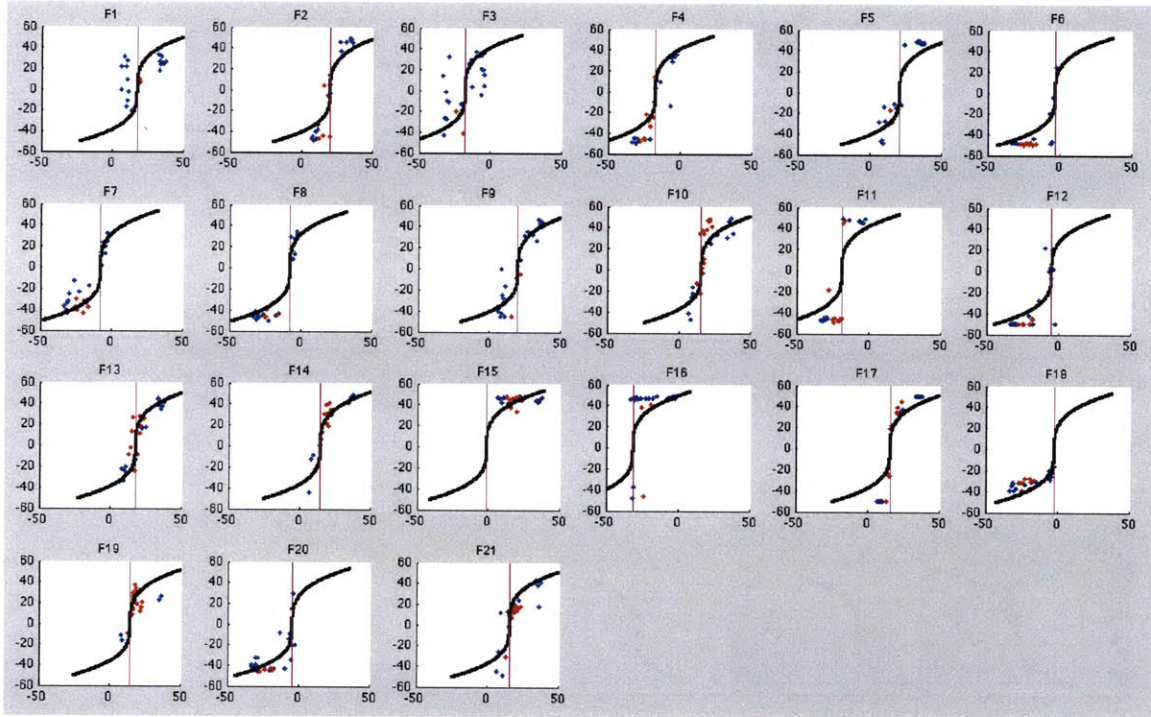
Figure A-9: Experienced-utility (EU) function with the *inferred* EU reference point (shown by the magenta line) for each respondent in the fear condition on Domain 2. The points show self-reported EUs: red diamond (safe option) and blue diamond (risky option). Note that the EU reference point is varied over respondents on Domain 2.
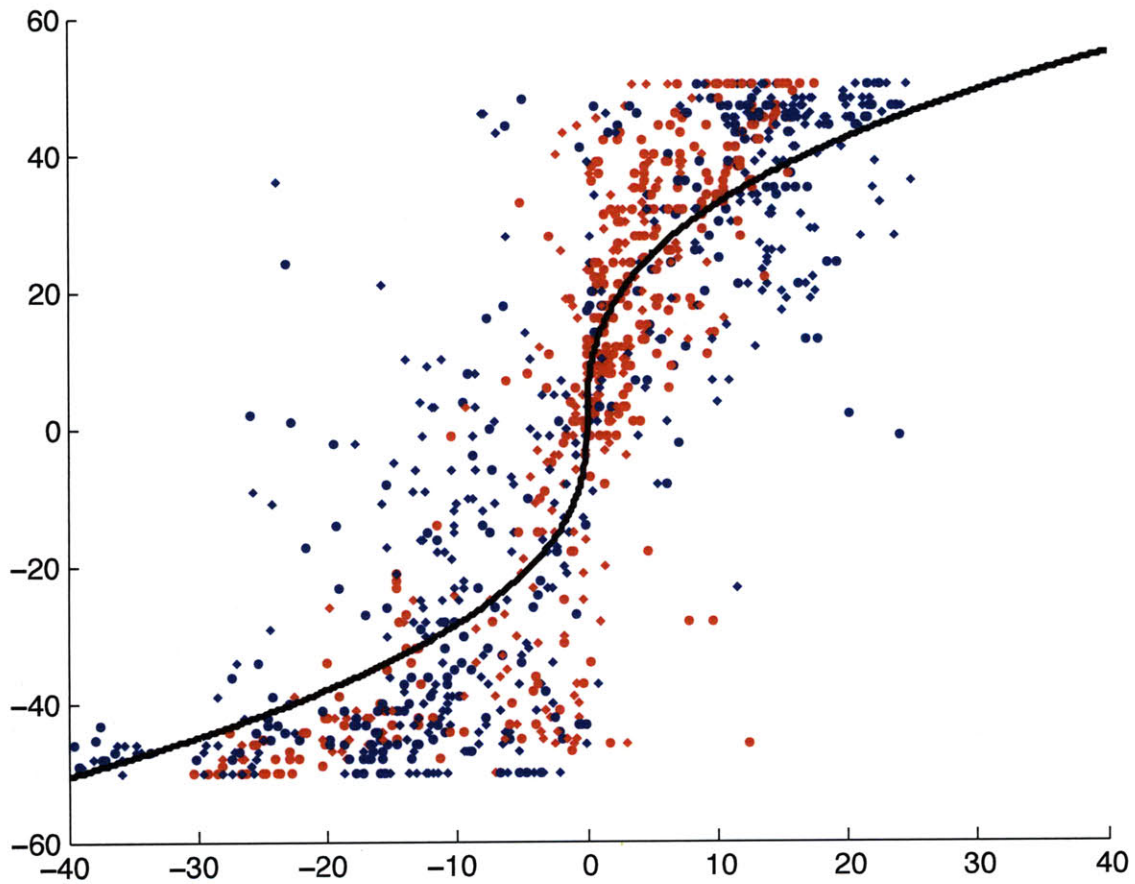
Figure A-10: Experienced-utility (EU) function in the economic fear condition (pooling over respondents). The points show self-reported EUs: red circle (safe option) and blue circle (risky option) on Domain 1, red diamond (safe option) and blue diamond (risky option) on Domain 2.
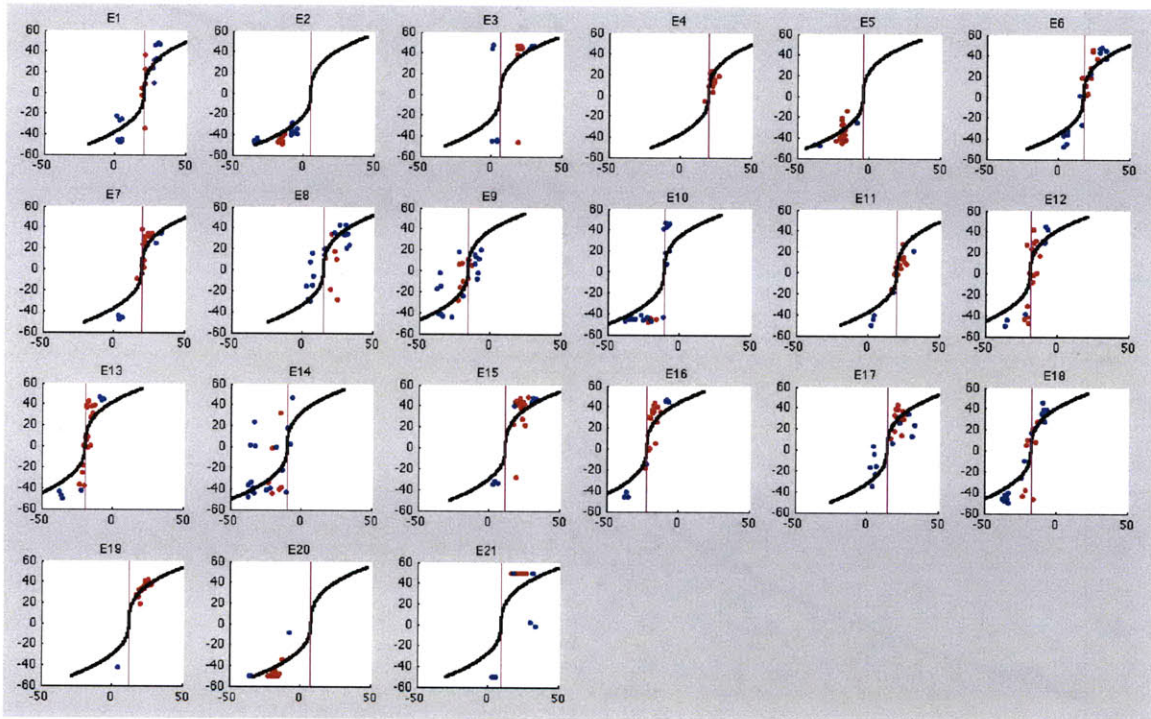
Figure A-11: Experienced-utility (EU) function with the *inferred* EU reference point (shown by the magenta line) for each respondent in the economic fear condition on Domain 1. The points show self-reported EUs: red circle (safe option) and blue circle (risky option). Note that the EU reference point is varied over respondents on Domain 1.
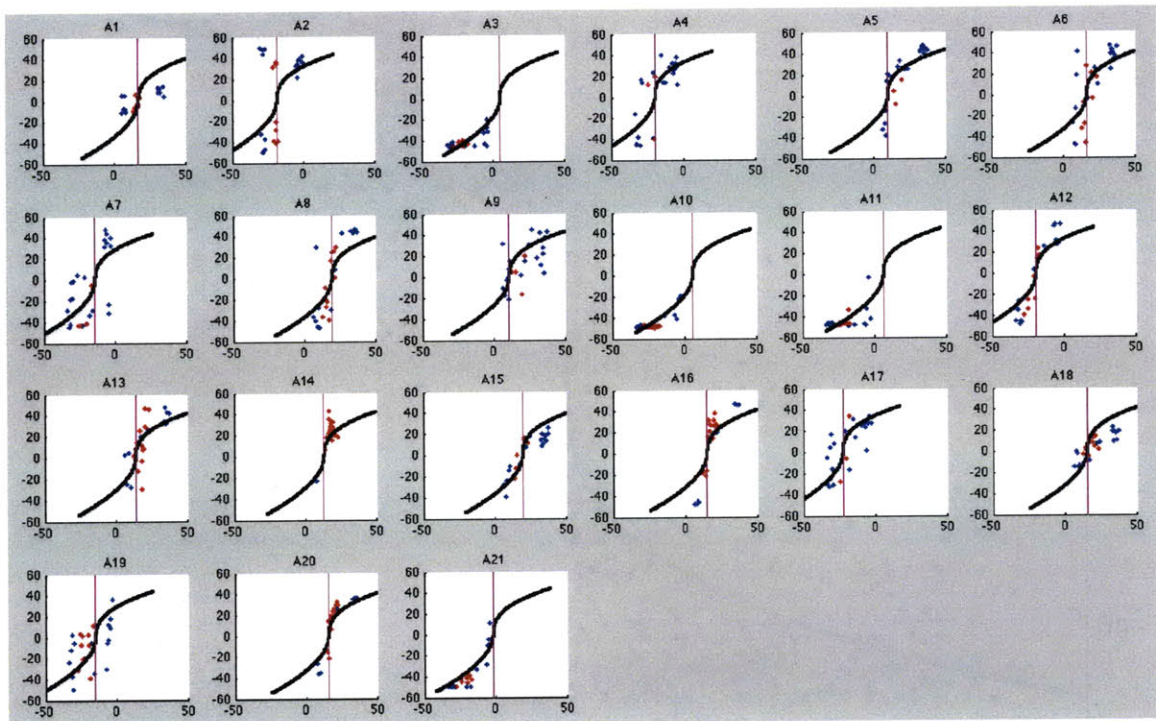
190
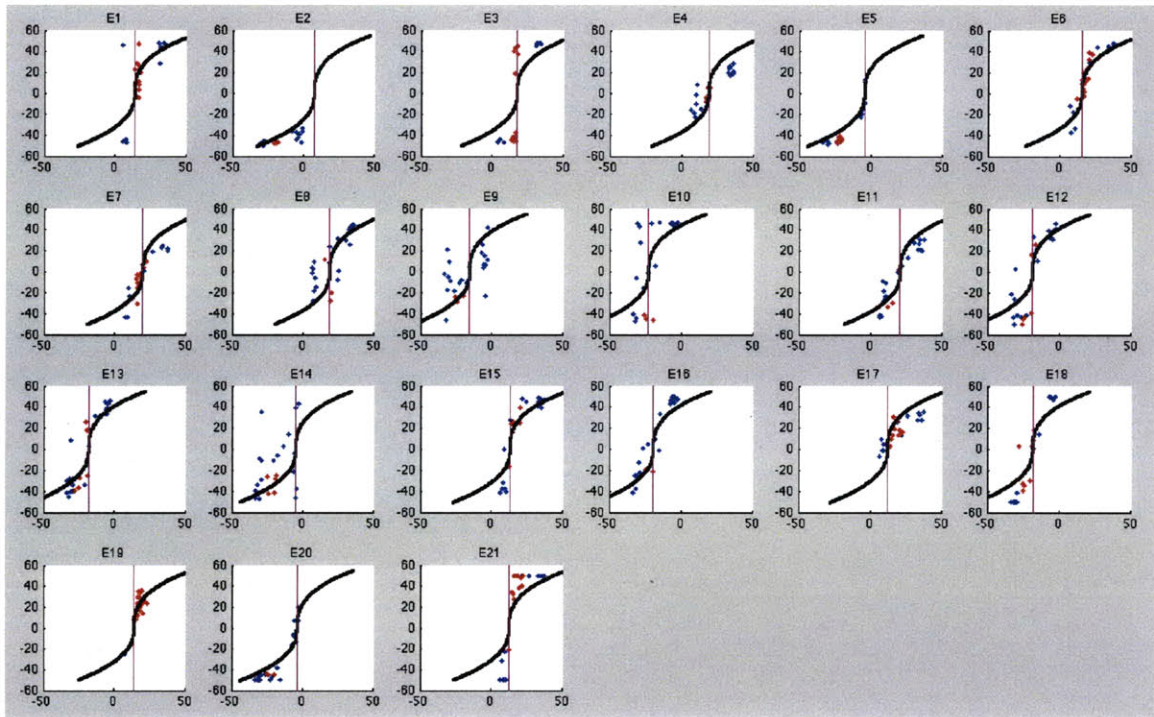
Figure A-12: Experienced-utility (EU) function with the *inferred* EU reference point (shown by the magenta line) for each respondent in the economic fear condition on Domain 2. The points show self-reported EUs: red diamond (safe option) and blue diamond (risky option). Note that the EU reference point is varied over respondents on Domain 2.

Appendix B

Inferring experienced-utility functions

Let $v(t)$ denote the respondents' self-reported *experienced utility* for the obtained outcome $x(t)$ on trial t during a task. Note that self-reported experienced-utility responses were converted to a scale $v(t)$ from -50 (=very displeased) to 50 (=very pleased). In each emotion condition, there were 21 respondents conducting two decision tasks (Domain 1 and Domain 2).

The following assumptions were made to infer the experienced-utility function from the experienced-utility responses of respondents in an emotion condition.

1. The self-reported experienced utility $v(t)$ on trial t during a task depends on the obtained outcome $x(t)$, the reference point x_{ref} and the shape parameters $(a, b, \lambda_G, \lambda_L)$ of the experienced-utility function

$$f_{EU}(x(t)|a, b, \lambda_G, \lambda_L, x_{ref}) = \begin{cases} \lambda_G(x(t) - x_{ref})^a, & x(t) \geq x_{ref} \\ -\lambda_L(x_{ref} - x(t))^b, & x(t) < x_{ref} \end{cases}$$

2. Respondents in the same emotion condition employ the common fixed shape parameters $(a, b, \lambda_G, \lambda_L)$ of the experienced-utility function, which were also fixed across tasks (Domain 1 and Domain 2).

3. The reference point x_{ref} is varied among respondents. Also, x_{ref} is fixed within a task, but varied between tasks.

For the mixed-effect model to infer the experienced-utility function of each emotion condition, the shape parameters were considered fixed effects and the reference-point parameters (of two domains) were random effects.

$x_{i,j}(t)$: obtained outcome on trial t during the task of domain $i(= 1,2)$ for respondent j

$v_{i,j}(t)$: self-reported experienced-utility on trial t during the task of domain i $(= 1,2)$ for respondent j $(-50 \leq v_{i,j}(t) \leq 50)$

$\hat{v}_{i,j}(t)$: estimated experienced-utility

$x_{ref\,i,j}$: reference point during the task of domain i $(= 1,2)$ for respondent j

$$\hat{v}_{i,j}(t) = f_{EU}(x_{i,j}(t)|a, b, \lambda_G, \lambda_L, x_{ref1,j}, x_{ref2,j}) = \begin{cases} \lambda_G(x_{i,j}(t) - x_{ref\,i,j})^a, & x_{i,j}(t) \geq x_{ref\,i,j} \\ -\lambda_L(x_{ref\,i,j} - x_{i,j}(t))^b, & x_{i,j}(t) < x_{ref\,i,j} \end{cases}$$

$$x_{i,j}(t) = x_{i,j}^{raw}(t) - m_{i,j}$$

$m_{i,j} =$ bet money (paid (+) / received (-)) during the task of domain i $(= 1,2)$ for respondent j

$$x_{ref1,j} = \bar{x}_{ref1} + \delta_{1,j}$$

$$x_{ref2,j} = \bar{x}_{ref2} + \delta_{2,j}$$

The model is:

    fixed effects $\beta = (a,\ b,\ \lambda_G,\ \lambda_L,\ \bar{x}_{ref1},\ \bar{x}_{ref2})$

random effects $\delta_j = (0, \ 0, \ 0, \ 0, \ \delta_{1,j}, \ \delta_{2,j})$

$$\varphi_j = \beta + \delta_j = (a, \ b, \ \lambda_G, \lambda_L, \ \bar{x}_{ref1} + \delta_{1,j}, \ \bar{x}_{ref2} + \delta_{2,j}) = (a, b, \lambda_G, \lambda_L, \ x_{ref1,j}, \ x_{ref2,j})$$

$$v_{i,j}(t) = \hat{v}_{i,j}(t) + \varepsilon_{i,j}(t) = f_{EU}(x_{i,j}(t)|\varphi_j) + \varepsilon_{i,j}(t)$$

$$\delta_j \sim N(0, \Psi) \text{ where } \Psi = diag(0, 0, 0, 0, var(\delta_{1,j}), var(\delta_{2,j}))$$

$$\varepsilon_{i,j}(t) \sim N(0, \sigma^2).$$

The model was fitted with data using NLMEFIT (nonlinear mixed-effects estimation) in the MATLAB Statistics Toolbox. NLMEFIT fits the model by maximizing an approximation to the marginal likelihood, i.e., with the random effects integrated out, and assumes that:

a) the random effects are multivariate normally distributed, and independent between groups, and

b) the observation errors are independent, identically normally distributed, and independent of the random effects.

By default, NLMEFIT fits a model where each model parameter is the sum of a corresponding fixed and random effect, and the covariance matrix of the random effects is diagonal, i.e., uncorrelated random effects.

Algorithm

In order to estimate the parameters of a nonlinear mixed effects model, we would like to choose the parameter values that maximize a likelihood function. These values are called the maximum likelihood estimates. The likelihood function can be written in the form

$$p(D|\beta, \sigma^2, \Psi) = \int p(D|\beta, \delta, \sigma^2) \, p(\delta|\Psi) \, d\delta$$

where

D is the response data,

β is the vector of population coefficients,

σ^2 is the residual variance,

Ψ is the covariance matrix for the random effects,

δ is the set of unobserved random effects.

Each $p()$ function on the right-hand-side is a normal (Gaussian) likelihood function that may depend on covariates.

Appendix C

Inferring predicted-utility functions

The self-reported *predicted utility* responses were linearly converted to a scale from -50 = very displeased to 50 = very pleased. Let $u(t|k)$ denote the converted measure of self-reported predicted utility response of option k (=1,2) on trial t during a task. Also, the self-reported confidence response on trial t was converted to a scale from -50 = not at all confident to 50 = very confident. Let $q(t)$ denote the converted measure of self-reported confidence response on trial t.

In each emotion condition, there were total 21 respondents conducting two decision tasks (Domain 1 and Domain 2).

The following model assumptions were made to infer the predicted-utility function from the predicted-utility responses of respondents in an emotion condition.

1. Prediction can be viewed as mapping the current-estimated outcome distribution into the predicted-utility distribution through the predicted-utility function whose shape depends on the current emotion state and confidence state. Note that the shape of the predicted-utility function in an experiment (emotion) condition is assumed to be trial-dependent due to the confidence state that may change over trials.

2. The self-reported predicted utility $u(t|k)$ of option k on trial t can be modeled by the average of the predicted-utility distribution.

3. The estimated outcome distribution of option k on trial t can be modeled by the set of previous outcomes obtained from selections of option k on previous trials $\{x(\tau)|c(\tau) = k, \tau = 1 \text{ to } t\}$, where $c(\tau)$ and $x(\tau)$ are the selected option and the obtained outcome from that selection on trial τ ($\leq t$).

4. The predicted-utility distribution of option k on trial t can be modeled by the predicted-utility samples $\{y(\tau)|c(\tau) = k, \tau = 1 \text{ to } t\}$ or the mapping of the set of previous outcomes of option k ($\{x(\tau)|c(\tau) = k, \tau = 1 \text{ to } t\}$) through the *current* shape of the predicted-utility function.

5. The predicted-utility sample $y(\tau)$ corresponding to $x(\tau)$ ($\tau=1$ to t) is computed using the predicted utility function:

$$f_{PU}(x(\tau)|a, b, \lambda_G(\tau), \lambda_L(\tau), x_{ref}) = y(\tau) = \begin{cases} \lambda_G(\tau)\,(x(\tau) - x_{ref})^a, & x(\tau) \geq x_{ref} \\ -\lambda_L(\tau)\,(x_{ref} - x(\tau))^b, & x(\tau) < x_{ref} \end{cases}$$

Note that sensitivity parameters $\lambda_G(t)$ and $\lambda_L(t)$ depend on the confidence measurement $q(t)$. Thus, $y(\tau)$ depends on the reference point x_{ref} and the shape parameters $(a, b, \lambda_G(t), \lambda_L(t))$ of the predicted utility function.

6. Sensitivity parameters $\lambda_G(t)$ and $\lambda_L(t)$ are functions of the confidence state variable $e(t)$ ranging from 0 (= not at all confident) to 1 (= very confident):

$e(t) = \frac{1}{1+\exp(-\kappa q(t))}$ where $\kappa = 0.1$

$\lambda_G(t) = \lambda_G(e(t)) = \lambda_{Gbase} + \lambda_{Gslope}(2\,e(t) - 1)$

$\lambda_L(t) = \lambda_L(e(t)) = \lambda_{Lbase} + \lambda_{Lslope}(2\,e(t) - 1)$

When $e(t) = 0.5$ (neutral confident), $\lambda_G(t) = \lambda_{Gbase}$ and $\lambda_L(t) = \lambda_{Lbase}$.

When $e(t) = 0$ (not at all confident), $\lambda_G(t) = \lambda_{Gbase} - \lambda_{Gslope}$ and $\lambda_L(t) = \lambda_{Lbase} - \lambda_{Lslope}$.

When $e(t) = 1$ (very confident), $\lambda_G(t) = \lambda_{Gbase} + \lambda_{Gslope}$ and $\lambda_L(t) = \lambda_{Lbase} + \lambda_{Lslope}$.

Also, $\lambda_G(t)$ and $\lambda_L(t)$ should be always positive. Note that λ_{Gbase} and λ_{Lbase} should be positive values, but that λ_{Gslope} and λ_{Lslope} may be either positive or negative.

7. The predicted utility of option k, denoted as $\hat{u}(t|k)$, is modeled by the average of

198

predicted-utility samples of option k (i.e., samples when $c(\tau) = k$ for $\tau = 1$ to t):
$\hat{u}(t|k) = E_{\tau=1:t}[y(\tau)|c(\tau) = k]$.

8. Respondents in the same emotion condition employ the common fixed shape parameters $(a, b, \lambda_{Gbase}, \lambda_{Lbase}, \lambda_{Gslope}, \lambda_{Lslope})$ of the predicted-utility function, which were also fixed across tasks (Domain 1 and Domain 2).

9. The reference point x_{ref} was varied among respondents. Also, x_{ref} was fixed within a task, but varied between tasks. Note that, for each task, it is assumed that the reference point of the predicted-utility function is independent of that of the experienced-utility function.

For the mixed-effect model to infer the predicted-utility function of each emotion condition, the shape parameters were considered fixed effects and the reference-point parameters (of two domains) were random effects.

$x_{i,j}(t)$: obtained outcome on trial t during the task of domain i ($= 1,2$) for respondent j

$c_{i,j}(t)$: option (1 or 2) selected on trial t during the task of domain i ($= 1,2$) for respondent j

$u_{i,j}(t|k)$: self-reported predicted-utility of option k ($= 1,2$) on trial t during the task of domain i ($= 1,2$) for respondent j ($-50 \leq u_{i,j}(t|k) \leq 50$)

$\hat{u}_{i,j}(t|k)$: model estimation of predicted-utility

$x_{ref\,i,j}$: reference point during the task of domain i ($= 1,2$) for respondent j

$y_{i,j}(t)$: predicted-utility sample corresponding to the obtained outcome $x_{i,j}(t)$

$$y_{i,j}(t) = f_{PU}(x_{i,j}(t)|a, b, \lambda_{Gbase}, \lambda_{Lbase}, \lambda_{Gslope}, \lambda_{Lslope}, x_{ref1,j}, x_{ref2,j})$$
$$= \begin{cases} \lambda_{G\,i,j}(t) \, (x_{i,j}(t) - x_{ref\,i,j})^a, & x_{i,j}(t) \geq x_{ref\,i,j} \\ -\lambda_{L\,i,j}(t) \, (x_{ref\,i,j} - x_{i,j}(t))^b, & x_{i,j}(t) < x_{ref\,i,j} \end{cases}$$

$e_{i,j}(t) = \frac{1}{1+\exp(-\kappa q_{i,j}(t))}$ where $\kappa = 0.1$ and $q_{i,j}(t) =$ confidence measurement

$\lambda_{G\,i,j}(t) = \lambda_{G\,i,j}(e_{i,j}(t)) = \lambda_{Gbase} + \lambda_{Gslope}(2\,e_{i,j}(t) - 1)$

$\lambda_{L\,i,j}(t) = \lambda_{L\,i,j}(e_{i,j}(t)) = \lambda_{Lbase} + \lambda_{Lslope}(2\,e_{i,j}(t) - 1)$

$\hat{u}_{i,j}(t|k) = E_{\tau=1:t}[y_{i,j}(\tau)|c(\tau) = k]$: average of predicted-utility samples of option k (i.e., samples when $c(\tau) = k$ for $\tau = 1$ to t)

$x_{i,j}(t) = x_{i,j}^{raw}(t) - m_{i,j}$

$m_{i,j} =$ bet money (paid (+) / received (-)) during the task of domain i (= 1,2) for respondent j

$x_{ref1,j} = \bar{x}_{ref1} + \delta_{1,j}$

$x_{ref2,j} = \bar{x}_{ref2} + \delta_{2,j}$

The model is:

    fixed effects $\beta = (a,\ b,\ \lambda_{Gbase},\ \lambda_{Lbase},\ \lambda_{Gslope},\ \lambda_{Lslope},\ \bar{x}_{ref1},\ \bar{x}_{ref2})$

    random effects $\delta_j = (0,\ 0,\ 0,\ 0,\ 0,\ 0,\ \delta_{1,j},\ \delta_{2,j})$

$\varphi_j = \beta + \delta_j = (a,\ b,\ \lambda_{Gbase},\ \lambda_{Lbase},\ \lambda_{Gslope},\ \lambda_{Lslope},\ \bar{x}_{ref1} + \delta_{1,j},\ \bar{x}_{ref2} + \delta_{2,j})$

    $= (a,\ b,\ \lambda_{Gbase},\ \lambda_{Lbase},\ \lambda_{Gslope},\ \lambda_{Lslope},\ x_{ref1,j},\ x_{ref2,j})$

$u_{i,j}(t|k) = \hat{u}_{i,j}(t|k) + \varepsilon_{i,j}(t|k) = E_{\tau=1:t}[y_{i,j}(\tau)|c(\tau) = k] + \varepsilon_{i,j}(t|k)$

    where $\hat{y}_{i,j}(t) = f_{PU}(x_{i,j}(t)|\varphi_j)$

$\delta_j \sim N(0, \Psi)$

$$\varepsilon_{i,j}(t|k) \sim N(0, \sigma^2).$$

Appendix D

Fixed sequences of random outcomes used in the experiments

* Gain frame on Domain 1:

Case 1: starting with a very low outcome on the first trial of the risky option

- Safe option: 19.85, 24.45, 18.20, 27.50, 19.01, 24.69, 17.13, 20.78, 25.32, 23.44, 22.47, 23.69, 22.49, 22.57, 19.44, 21.93, 22.43, 21.77, 23.57, 23.76, 22.49, 21.13, 20.86, 20.77, 22.97, 20.96, 20.99, 22.61, 19.82, 20.94

- Risky option: 3.99, 32.01, 4.98, 31.41, 18.00, 33.41, 2.63, 30.18, 6.46, 13.44, 30.64, 29.27, 5.65, 20.61, 29.08, 4.76, 28.80, 9.56, 33.43, 7.24, 14.92, 31.26, 21.44, 5.40, 31.15, 6.10, 16.25, 31.14, 28.02, 4.45

Case 2: starting with a very high outcome on the first trial of the risky option

- Safe option: 24.15, 19.55, 25.80, 16.50, 24.99, 19.31, 26.87, 23.22, 18.68, 20.56, 21.53, 20.31, 21.51, 21.43, 24.56, 22.07, 21.57, 22.23, 20.43, 20.24, 21.51, 22.87, 23.14, 23.23, 21.03, 23.04, 23.01, 21.39, 24.18, 23.06

- Risky option: 32.01, 3.99, 31.02, 4.59, 18.00, 2.59, 33.37, 5.82, 29.54, 22.56, 5.36, 6.73, 30.35, 15.39, 6.92, 31.24, 7.20, 26.44, 2.57, 28.76, 21.08, 4.74, 14.56, 30.60, 4.85, 29.90, 19.75, 4.86, 7.98, 31.55

* Loss frame on Domain 1:

Case 1: starting with a very low outcome on the first trial of the risky option

- Safe option: -20.15, -15.55, -21.80, -12.50, -20.99, -15.31, -22.87, -19.22, -14.68, -16.56, -17.53, -16.31, -17.51, -17.43, -20.56, -18.07, -17.57, -18.23, -16.43, -16.24, -17.51, -18.87, -19.14, -19.23, -17.03, -19.04, -19.01, -17.39, -20.18, -19.06

- Risky option: -36.01, -7.99, -35.02, -8.59, -22.00, -6.59, -37.37, -9.82, -33.54, -26.56, -9.36, -10.73, -34.35, -19.39, -10.92, -35.24, -11.20, -30.44, -6.57, -32.76, -25.08, -8.74, -18.56, -34.60, -8.85, -33.90, -23.75, -8.86, -11.98, -35.55

Case 2: starting with a very high outcome on the first trial of the risky option

- Safe option: -15.85, -20.45, -14.20, -23.50, -15.01, -20.69, -13.13, -16.78, -21.32, -19.44, -18.47, -19.69, -18.49, -18.57, -15.44, -17.93, -18.43, -17.77, -19.57, -19.76, -18.49, -17.13, -16.86, -16.77, -18.97, -16.96, -16.99, -18.61, -15.82, -16.94

- Risky option: -7.99, -36.01, -8.98, -35.41, -22.00, -37.41, -6.63, -34.18, -10.46, -17.44, -34.64, -33.27, -9.65, -24.61, -33.08, -8.76, -32.80, -13.56, -37.43, -11.24, -18.92, -35.26, -25.44, -9.40, -35.15, -10.10, -20.25, -35.14, -32.02, -8.45

* Gain frame on Domain 2:

Case 1: starting with a very low outcome on the first trial of the risky option

- Safe option: 15.85, 20.45, 14.20, 23.50, 15.01, 20.69, 13.13, 16.78, 21.32, 19.44, 18.47, 19.69, 18.49, 18.57, 15.44, 17.93, 18.43, 17.77, 19.57, 19.76, 18.49, 17.13, 16.86, 16.77, 18.97, 16.96, 16.99, 18.61, 15.82, 16.94

- Risky option: 7.99, 36.01, 8.98, 35.41, 22.00, 37.41, 6.63, 34.18, 10.46, 17.44, 34.64, 33.27, 9.65, 24.61, 33.08, 8.76, 32.80, 13.56, 37.43, 11.24, 18.92, 35.26, 25.44, 9.40, 35.15, 10.10, 20.25, 35.14, 32.02, 8.45

Case 2: starting with a very high outcome on the first trial of the risky option

- Safe option: 20.15, 15.55, 21.80, 12.50, 20.99, 15.31, 22.87, 19.22, 14.68, 16.56, 17.53, 16.31, 17.51, 17.43, 20.56, 18.07, 17.57, 18.23, 16.43, 16.24, 17.51, 18.87, 19.14, 19.23, 17.03, 19.04, 19.01, 17.39, 20.18, 19.06

- Risky option: 36.01, 7.99, 35.02, 8.59, 22.00, 6.59, 37.37, 9.82, 33.54, 26.56, 9.36, 10.73, 34.35, 19.39, 10.92, 35.24, 11.20, 30.44, 6.57, 32.76, 25.08, 8.74, 18.56, 34.60, 8.85, 33.90, 23.75, 8.86, 11.98, 35.55

* Loss frame on Domain 2:

Case 1: starting with a very low outcome on the first trial of the risky option

- Safe option: -24.15, -19.55, -25.80, -16.50, -24.99, -19.31, -26.87, -23.22, -18.68, -20.56, -21.53, -20.31, -21.51, -21.43, -24.56, -22.07, -21.57, -22.23, -20.43, -20.24, -21.51, -22.87, -23.14, -23.23, -21.03, -23.04, -23.01, -21.39, -24.18, -23.06

- Risky option: -32.01, -3.99, -31.02, -4.59, -18.00, -2.59, -33.37, -5.82, -29.54, -22.56, -5.36, -6.73, -30.35, -15.39, -6.92, -31.24, -7.20, -26.44, -2.57, -28.76, -21.08, -4.74, -14.56, -30.60, -4.85, -29.90, -19.75, -4.86, -7.98, -31.55

Case 2: starting with a very high outcome on the first trial of the risky option

- Safe option: -19.85, -24.45, -18.20, -27.50, -19.01, -24.69, -17.13, -20.78, -25.32, -23.44, -22.47, -23.69, -22.49, -22.57, -19.44, -21.93, -22.43, -21.77, -23.57, -23.76, -22.49, -21.13, -20.86, -20.77, -22.97, -20.96, -20.99, -22.61, -19.82, -20.94

- Risky option: -3.99, -32.01, -4.98, -31.41, -18.00, -33.41, -2.63, -30.18, -6.46, -13.44, -30.64, -29.27, -5.65, -20.61, -29.08, -4.76, -28.80, -9.56, -33.43, -7.24, -14.92, -31.26, -21.44, -5.40, -31.15, -6.10, -16.25, -31.14, -28.02, -4.45

Bibliography

[1] B. W. Balleine and A. Dickinson. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*, 37(4-5):407–19, 1998.

[2] N. Barberis, M. Huang, and T. Santos. Prospect theory and asset prices. *Quarterly Journal of Economics*, 116(1):1–53, 2001.

[3] A. Barto and O. Simsek. Intrinsic motivation for reinforcement learning systems. *Proceedings of the Thirteenth Yale Workshop on Adaptive and Learning Systems, New Haven, CT, Yale University*, 2005.

[4] A. Bechara and A. R. Damasio. The somatic marker hypothesis: a neural theory of economic decision. *Games and Economic Behavior*, 52(2):336–372, 2005.

[5] A. Bechara, H. Damasio, D. Tranel, and A. Damasio. Deciding advantageously before knowing the advantageous strategy. *Science*, 275:1293–1295, 1997.

[6] KC. Berridge and TE. Robinson. Parsing reward. *Trends in Neurosciences*, 26(9), 2003.

[7] Kent C. Berridge and J. Wayne Aldridge. Decision utility, incentive salience, and cue-triggered 'wanting'. *To appear in Psychology of Action, 2nd Edition. Bargh, Gollwitzer, & Morsella (Eds.), Oxford University Press*, 2006.

[8] J. R. Busemeyer and J. T. Townsend. Decision field theory: A dynamic cognition approach to decision making in an uncertain environment. *Psychological Review*, 100(3):432–459, 1993.

[9] C. Camerer and T. H. Ho. Experience-weighted attraction learning in normal form games. *Econometrica*, 67(4):827–874, 1999.

[10] G.L. Clore, K. Gasper, and E. Garvin. Affect as information. *Handbook of affect and social cognition*, pages 121–144, 2001.

[11] J. Cohen. Statistical power analysis. *Current Directions in Psychological Science*, 1(3):98–101, 1992.

[12] J. B. Cohen, M. T. Pham, and E. B. Andrade. The nature and role of affect in consumer behavior. *Handbook of Consumer Psychology, Curtis P. Haugtvedt, Paul Herr, and Frank Kardes (Eds.)*, 2006.

[13] J.D. Cohen, S.M. McClure, and A.J. Yu. Should i stay or should i go? how the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1481):933–942, 2007.

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

[15] N. D. Daw, J. P. O'Doherty, P. Dayan, B. Seymour, and R. J. Dolan. Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095):876–9, 2006.

[16] P. Dayan, Y. Niv, B. Seymour, and N. D. Daw. The misbehavior of value and the discipline of the will. *Neural Networks*, 19:1153–1160, 2006.

[17] R. Dearden, N. Friedman, and S. Russell. Bayesian q-learning. *Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI-98)*, pages 761–768, 1998.

[18] R.O. Duda, P.E. Hart, and D.G. Stork. *Pattern classification.* Citeseer, 2001.

[19] I. Erev and G. Barron. On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review*, 112(4):912–931, 2005.

[20] C. R. Fox and L. Hadar. Decisions from experience= sampling error+ prospect theory: Reconsidering hertwig, barron, weber & erev (2004). *Judgment and Decision Making*, 1(2):159–161, 2006.

[21] P. W. Glimcher and A. Rustichini. Neuroeconomics: The consilience of brain and decision. *Science*, 306(5695):447, 2004.

[22] R. Hau, T. J. Pleskac, J. Kiefer, and R. Hertwig. The description-experience gap in risky choice: the role of sample size and experienced probabilities. *Journal of Behavioral Decision Making*, 21(5), 2008.

[23] R. Hertwig, G. Barron, E. U. Weber, and I. Erev. Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, 15:534–539, 2004.

[24] C. K. Hsee and R. Hastie. Decision and experience: why don't we choose what makes us happy? *TRENDS in Cognitive Sciences*, 10(1):32, 2006.

[25] A. M. Isen. An influence of positive affect on decision making in complex situations: Theoretical issues with practical implications. *Journal of Consumer Psychology*, 11(2):75–85, 2001.

[26] A. M. Isen, T. E. Nygren, and F. G. Ashby. Influence of positive affect on the subjective utility of gains and losses: It is just not worth the risk. *Journal of Personality and Social Psychology*, 55(5):710–17, 1988.

[27] L. P. Kaelbling. *Learning in Embedded Systems*. MIT Press, 1993.

[28] D. Kahneman. Experienced utility and objective happiness: A moment-based approach. *Choices, values, and frames*, pages 673–692, 2000.

[29] D. Kahneman. Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review*, 93(5):1449–1475, 2003.

[30] D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292, 1979.

[31] Daniel Kahneman, Peter P. Wakker, and Rakesh Sarin. Back to bentham? explorations of experienced utility. *The Quarterly Journal of Economics*, 112(2):375–405, 1997.

[32] J. S. Lerner. Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition & Emotion*, 14(4):473–493, 2000.

[33] J. S. Lerner, R. M. Gonzalez, D. A. Small, and B. Fischhoff. Effects of fear and anger on perceived risks of terrorism: A national field experiment. *Psychological Science*, 14(2):144–150, 2003.

[34] J. S. Lerner and D. Keltner. Fear, anger, and risk. *Journal of Personality and Social Psychology*, 81(1):146–159, 2001.

[35] J. S. Lerner, D. A. Small, and G. Loewenstein. Heart strings and purse strings. *Psychological Science*, 15(5):337–341, 2004.

[36] G. Loewenstein and J. S. Lerner. The role of affect in decision making. *Handbook of affective science*, pages 619–642, 2003.

[37] G. Loewenstein and T. O'Donoghue. Animal spirits: Affective and deliberative processes in economic behavior. 2005.

[38] G. Loewenstein, E. U. Weber, C. K. Hsee, and N. Welch. Risk as feelings. *Psychological Bulletin*, 127(2):267–286, 2001.

[39] N. Meuleau and P. Bourgine. Exploration of multi-states environments: Local measures and back-propogation of uncertainty. *Machine Learning*, 1998.

[40] N. Naqvi, B. Shiv, and A. Bechara. The role of emotion in decision making: A cognitive neuroscience perspective. *Current Directions in Psychological Science*, 15(5):260–264, 2006.

[41] P. M. Niedenthal. Embodying emotion. *Science*, 316(5827):1002, 2007.

[42] J. Panksepp. *Affective Neuroscience*. Oxford University Press, 1998.

[43] M. T. Pham. The logic of feeling. *Journal of Consumer Psychology*, 14(4):360–369, 2004.

[44] M. T. Pham. Emotion and rationality: A critical review and interpretation of empirical evidence. *Review of General Psychology*, 11(2):155–178, 2007.

[45] M.Z. Poh, N.C. Swenson, and R.W. Picard. A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. *IEEE Transactions on Biomedical Engineering*, 57(5):1243–1252, 2010.

[46] R. Raghunathan and M. T. Pham. All negative moods are not equal: Motivational influences of anxiety and sadness on decision making. *Organizational Behavior and Human Decision Processes*, 79(1):56–77, 1999.

[47] R. Raghunathan, M. T. Pham, and K. P. Corfman. Informational properties of anxiety and sadness, and displaced coping. *Journal of Consumer Research*, 32, 2006.

[48] L. A. Real. Animal choice behavior and the evolution of cognitive architecture. *Science*, 253(5023):980–986, 1991.

[49] J. Rottenberg, RR Ray, and J.J. Gross. Emotion elicitation using films. *The handbook of emotion elicitation and assessment*, pages 9–28, 2007.

[50] A. G. Sanfey, G. Loewenstein, S. M. McClure, and J. D. Cohen. Neuroeconomics: cross-currents in research on decision-making. *Trends in cognitive sciences*, 10(3), 2006.

[51] N. Schwarz and G.L. Clore. How do I feel about it? The informative function of affective states. *Affect, cognition, and social behavior*, pages 44–62, 1988.

[52] S. Shafir, T. Reich, E. Tsur, I. Erev, and A. Lotem. Perceptual accuracy and conflicting effects of certainty on risk-taking behaviour. *Nature*, 453(7197):917–920, 2008.

[53] B. Shiv and A. Fedorikhin. Heart and mind in conflict: The interplay of affect and cognition in consumer decision making. *The Journal of Consumer Research*, 26(3):278–292, 1999.

[54] B. Shiv, G. Loewenstein, A. Bechara, H. Damasio, and A. R. Damasio. Investment behavior and the negative side of emotion. *Psychological Science*, 16(6):435, 2005.

[55] O. Simsek and A. Barto. An intrinsic reward mechanism for efficient exploration. *Proceedings of the 23rd international conference on Machine learning*, 2006.

[56] S. Singh, A. G. Barto, and N. Chentanez. Intrinsically motivated reinforcement learning. *Advances in Neural Information Processing*, 18, 2004.

[57] A. Stout, G. Konidaris, and A. Barto. Intrinsically motivated reinforcement learning: A promising framework for developmental robot learning. *AAAI Spring Symposium on Developmental Robotics, Stanford, California, March*, 2005.

[58] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 1998.

[59] R. S. Sutton, D. Precup, and S. Singh. Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence*, 112:181–211, 1999.

[60] L. L. Thurstone. Psychophysical analysis. *The American Journal of Psychology*, 38(3):368–389, 1927.

[61] A. Tversky and D. Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323, 1992.

[62] T.D. Wickens. *Elementary signal detection theory*. Oxford University Press, USA, 2002.

[63] E. Yechiam, J.R. Busemeyer, J.C. Stout, and A. Bechara. Using cognitive models to map relations between neuropsychological disorders and human decision-making deficits. *Psychological Science*, 16(12):973, 2005.