Emotions in Action

An inquiry into the explanation of decision-making in the real economic world

Emoties in Actie

Op zoek naar de verklaring van besluitvorming in the échte economische wereld

Thesis

To obtain the degree of Doctor from the
Erasmus University Rotterdam
By the command of the rector magnificus

prof.dr. S.W.J. Lamberts

And in accordance with the decision of the Doctorate Board
The public defense shall be held on
16 November **2006 at 09:00hrs**

By Roberta Muramatsu,
born in São Caetano do Sul, São Paulo, Brazil

Doctoral Committee

Promotor:
Prof.dr. I.U. Mäki

Other Members:
Prof.dr. M. Blaug
Prof.dr. A. Klamer
Prof.dr. W.F. van Raaij

Copromotor:
Dr. J.J. Vromen

To think is an act. To feel is a fact.

The Hour of Star, Clarice Lispector

Emotions are mechanisms that set the brain's highest goals. Once triggered by a propitious moment, an emotion triggers the cascade of subgoals and sub-subgoals that we call thinking and acting…no sharp line divides thinking from feeling.

How the Mind Works, Steven Pinker

To Bela, Toninho, Renata, Rita, Paulo, Tutti, Ursula and Zulu - the objects of my wisest emotions

Contents

Preface and Acknowledgements

This book is a product of long years of reading, writing and re-writing. This was due to the complexity of my chosen research topic – scientific explanation of decision-making behavior and the role emotions play in it. I am very grateful to Jack Vromen and Uskali Mäki for helping me to undertake the most challenging intellectual task of my life. The focus of my dissertation has changed substantively from its initial project. Without their help, it would have been very difficult to structure my ideas and select an effective argumentation strategy.

I owe very special thanks to Jack Vromen. He did devote a lot of time and energy to read and comment (in person and through hundreds of emails) on all drafts of the book (and PhD research proposals). I would like to say 'muitissimo obrigada' for having guided me to navigate in a not calm sea of ideas, concepts, methodologies and models. Under many stages of the research, I was about to get drowned by so many conceptual, theoretical and empirical puzzles. But Jack was always ready to help me out. He taught me one very important lesson that stubborn research-driven students (like me) cannot understand very easily- sometimes less is more. As a university lecturer, I am trying to do the same with my young ambitious undergraduate students.

I am very happy that I had the chance to have Uskali Mäki as my supervisor. In the last six years I learned a lot from him. Without his careful reading, advice and commentaries on my PhD dissertation, I am not sure I could have overcome many serious challenges posed by my ambitious project of interdisciplinary research. Interacting with Uskali made me learn and value the fascinating move from bold claims and statement to defensible arguments.

My EIPEan friend Caterina Marchionni also contributed a lot to my PhD research and she also deserves many thanks. Like my supervisors, she took the trouble of reading carefully the long chapters of my thesis. Cate helped me to detect and to revise important (hidden) claims and premises of my thesis.

Many other people outside EIPE inspired me and helped me to think over my earlier drafts of this book, presented in conferences and workshops. I would like to reveal my gratefulness to Ana Maria Bianchi, Yaniv Hanoch, Ramón Fernandez, Laura Valadão, and Eduardo Giannetti for their very insightful comments and suggestions.

Living in the Netherlands was a big challenge to a Brazilian in various respects (cultural, social, intellectual, etc). But all the initial difficulties were quickly overcome because I was too eager to learn from interaction with the nice teachers and PhD students I met in EIPE. Attending their courses, mini-courses and seminars were very much inspiring and gave me much food for thought. I would like to thank all EIPEans for having created such an exciting intellectual environment. My life in Rotterdam gained very nice colors and brightness and I owe this to my family. Many thanks for the (moral and sometimes financial) support that I received from them anytime I needed most. Despite the geographical distance (a whole ocean putting us apart), they were always there for me. Without their love, support and patience I could not become the person I am today. I would like to express my love and gratitude by dedicating this book to them.

I think I got easily adapted to the Dutch way of living – going shopping (*boodschappen doen*) in the street market, ridding a bike (*fietsen*) to go everywhere – due to the special friends I made there. Many of them were foreigners who struggled to make a gratifying and worth-living life abroad. Some others were Dutch people fascinated by cultural diversity and its challenges. I would like to thank Caterina Marchionni, Ana Cordeiro dos Santos, Gülbahar Tezel, Susana Graça, Altug Yalcintas, Alessandro Lanteri, Estefano Ficco, Vilma Moraes, Marcio Fugiy, Ayako Fugiy, Guilherme Fugiy, Henrique Fugiy, Ger van Werkhoven, Tony Lelieveldt, Hermann Lelieveldt, among others for their companionship, support, trust and understanding.

I feel very much indebted to CAPES foundation because it gave me the opportunity to do my doctorate research in the Netherlands. I am particularly grateful to Sandra Hugo de Jesus, Gérlia Chaves and Angélica Silva. They were always willing to help to deal with

formal matters concerning my PhD research schedule and process. I would like to thank Tony Lelieveldt and Michiel Wielema for taking the trouble of translating my thesis' summary into Dutch. Finally, I need to express my gratitude to Mark Dambrink, Lizzy Patilaya, Linda Degener and Professor Dr. L. van Bunge. Without their help and support, it would have been quite difficult for me to arrange many practical things for the most important ritual in my intellectual career - the defense of my PhD thesis.

ABSTRACT

Decision-making in the real world of boundedly rational agents is a complex phenomenon. The large literature on anomalies reflects doubts about whether the resemblance between the conventional economic choice model and reality is sufficiently close. In response, many have engaged in reformist theoretical modeling strategies. These strategies are here analyzed in terms of their potential contribution to explanatory progress. The notion of explanatory progress is understood in terms of expanding the scope of explanations and penetrating deeper into the causal mechanisms of behavior. The question is whether invoking emotions and brain processes as explanatory factors will promote progress.

Based on theoretical advancements within the study of intertemporal choice, decision under risk and prosocial choice, it is argued that scope expansion will be attained through the development of models that incorporate effective emotional processes. Consulting recent work in the neurosciences and their applications, underlying causal mechanisms are sought by sketching a theoretical model of choice that combines the levels of mind and brain.

CHAPTER 1
WHAT ARE THE ISSUES?

> The basic difficulty in economic theory is the philosophical problem of the meaning of explanation in connection with human behavior.
>
> Frank Knight, "Economic Psychology and the Value Problem", *Quarterly Journal of Economics*, 1925, p. 372

1. Introduction

Decision-making is a rather complex and pervasive phenomenon and therefore has prompted a great deal of research in economics, psychology, and more recently, in the neurosciences. Many departments of human affairs can be re-described in terms of specialized decision problems, or choice tasks.[1] Yet the paradigmatic choice framework (i.e. subjective expected utility theory) does not seem to offer satisfactory accounts of choice behavior. There is systematic empirical evidence that actual people's judgments and decisions often deviate from the predictions made by standard decision theory.

When recurrent and systematic, findings of this kind challenge the ability of the conventional approach to truthfully represent relevant aspects of mental and social reality given its scientific purposes of prediction and explanation. These troubling behavior patterns are often called choice anomalies, puzzles, or paradoxes. The current work assumes that economically significant 'choice anomalies' serve as triggers for theoretical innovations in search for explanatory improvements.

This book is a study of the task of explaining decision-making behavior in the real economic world. The suggestion here is that a systematic investigation of some methodological issues of economic analysis helps us to work toward more thorough and

1 Lionel Robbins famously characterized economics as "the science, which studies behavior as a relationship between ends and scarce means which have alternative uses' (1935/1962, p.16).

far-reaching accounts of choice behavior. The point of departure is the vision that a satisfactory explanation of decision-making behavior should uncover significant processes or mechanisms whereby such an *explanandum* phenomenon comes about. In so doing, an explanatory account offers scientific understanding by showing how the phenomenon participates in a complex world of causal and other dependencies, often by tracing the (causal) chain of events that give rise to manifest choice behavior. The present work makes seven main claims, denoted by C's:

(C1) Empirical discovery of anomalies prompt and motivate reformist theoretical modeling efforts. When recurrent and systematic, anomalies indicate that the standard model of choice failed to meet its alleged explanatory tasks.

(C2) Model revision or modification involves relaxation of idealizing and simplifying assumptions, which can be interpreted as supplementation (de-isolation) and replacement (re-isolation) of the previously selected set of explanatory factors.

(C3) Explanation (prediction) of choice anomalies and identification of (causally) significant processes or mechanisms for behavior are the two main goals that drive model building and subsequent revision by behavioral researchers.

(C4) Behavioral models of choice that uncover major processes in the production of manifest behavior might yield improved understanding of why, how and under what conditions certain behavior patterns occur.

(C5) Behavioral models that incorporate the non-negligible roles that emotions play in decision-making behavior contribute to progress as scope expansion. They do so because they can accommodate behaviors already covered by standard choice theory, as well as anomalies and new explanandum phenomena within the domains of intertemporal choice, decision under risk and prosocial choice.

(C6) A theoretical model of decision-making at the levels of mind and brain might attain explanatory progress as enhanced causal articulation since it unveils a complex chain of emotional and higher order cognitive processes and states that might produce economically relevant patterns of behavior.

(C7) A two-level account of decision-making might also bring explanatory progress as causal penetration because it uncovers those neural structures and information processing activities that physically execute those causally significant mental processes for behavior.

The main contribution of this monograph is to provide a philosophical analysis of the complex issue of explanation of actual decision-making behavior. It can be interpreted as a response to one behavioral economist's vision that 'as a rule, it is bad to spend time on methodological and broad stroke issues rather than the nitty-gritty of the phenomena being studied (Rabin 2002, p. 659). The present work is an outcome of having spent quite some time on just such issues. However, instead of adopting those prepackaged criteria that characterizes an 'off-the-shelf philosophy of science approach', I propose a methodological analysis that relies on the actual practice of behavioral economists and its psychologists' and neuroscientists neighbors. The suggested philosophical inquiry consists in presenting and scrutinizing conceptual and metatheoretical ideas that give rise to advances in accounts of choice behavior that take emotions (among higher-order cognition) seriously. Another related contribution is to tentatively sketch a theoretical model that gives a mechanistic explanation of decision-making and the roles emotions play in it. In so doing, this work aims to develop an argument that follows up on Simon's explanatory account of choices in the real world of boundedly rational agents.

2. The growth of psychology-informed economic explanation: a move towards explanatory progress

Contemporary behavioral or psychological explanations gained momentum as recognized research tasks in economics in the 1990s when behavioral economists headed departments in leading universities and earned prestigious awards such as the Nobel Prize or the John Bates Clark medal (Loewenstein 1999, Uchitelle 2001, Rabin 2002). Journals

have published many articles by behavioral economists and there has been a great deal of theoretical and experimental research based on a psychological economic framework.

This monograph provides an interpretation of the revival of behavioral accounts in the final decades of the 20th century. It puts forth the idea that two ambitions - prediction of anomalies and description of hidden processes or mechanisms for choice behavior (anomalies included) - motivate economists to question their models' psychological assumptions and to reform those that are believed to exclude explanatorily relevant factors.

Most economists did not take very seriously severe criticisms of conventional decision theory (and its Olympian vision of rational choice) by early behavioral economists of the 20th century (e.g. Katona 1951 and 1980, Edwards 1954, Simon 1955, 1956 and 1983). This was largely due to the popularity of ideas that were to receive a persuasive expression in Milton Friedman's methodological argument that expected utility theory meets the purposes of economic accounts, i.e. to make accurate predictions of a wide range of choice behaviors at the market level, despite of its unrealistic assumptions (and its silence in relation to processes or mechanisms of actual behavior).

The present monograph is concerned with metatheoretical issues and convictions that guide ongoing modifications of idealizing and simplifying behavioral assumptions of economic accounts. For that reason, it devotes some time to discuss contemporary economists' vision that "at the core of behavioral economics is the conviction that increasing the realism of psychological underpinnings of economic analysis will improve the field on its own terns" (Camerer and Loewenstein 2004, p.3). Driven by a realist perspective on theorizing and explanation, I propose some qualifications to the project of making psychological assumptions less unrealistic and the consequences of this for the goal of explanatory progress. Furthermore, I diagnose two theorizing strategies that seem to guide psychological economic explanations. The first is called the incremental strategy. It does not radically depart from the neoclassical approach (and its underlying explanatory scheme). Rather this reformist strategy consists in relaxing specific

behavioral assumptions for the sake of predicting certain anomalies.[2] More recently, there has been a second strategy, which I dub the process-description approach. It deviates from the mainstream utility approach (and its underlying vision of positive economic theorizing) in that it gives rise to accounts that describe significant processes by which actual behavior is produced (Camerer, Loewenstein and Prelec 2005).

On the basis of (empirical) behavioral models that rely on an interdisciplinary research strategy I also sketch a theoretical model of choice behavior that draws on Simon's account of boundedly rational decisions and on recent findings in the brain sciences and evolutionary psychology. The proposed model offers a description of the machinery of human decision making at mind and physical brain levels of analysis. It purports to uncover the important roles that emotions play in the (causal) production of actual choices within three puzzling domains of economic analysis – intertemporal choice, decision under risk and uncertainty, and prosocial choice.

3. Emotions within the (economic) explanation of choice

Many of us have the impression that emotions play an important role in a commonsense explanation of decision-making. Yet most practicing economists do not incorporate emotions explicitly into their formal accounts of choice behavior.

If this is so, we are tempted to suggest that "economists have totally neglected the most important aspect of their subject matter" (Elster 1996, p.1386).

Economic models often assume that emotions are sources of preferences (arguments of a utility function, for instance) or constraint parameters (i.e. emotional costs or benefits associated with particular courses of actions).[3] This is partly so due to a long-standing tradition of viewing emotions as 'sand' clogging the machinery of rational choice.

2 In Rabin's (2002) words,

> "This research program is not only built upon the premise that mainstream economic methods are great, but also that mainstream economic assumptions are great. It does not abandon the correct insights from neoclassical economics, but supplements these insights with the insights to be had from realistic new assumptions." (pp. 658-659)

3 For instance, see Loomes and Sudgen (1982), Hirschleifer (1987), Frank (1988) and Becker (1996).

Economics is often regarded as a discipline concerned with rational choice explanations.[4] The present work investigates why and how emotions have been brought into the toolbox of economic analysis. It claims that the difficulties in predicting anomalies and in detecting processes or mechanisms for behavior serve as motives for elaborating on accounts that take emotions seriously.

In this book, I advance the thesis that, under some conditions, emotions and feelings assist (rather than undermine) quick and effective inferences and choices. My argument is built on Simon's explanatory account of boundedly rational choice. To him, "human rational behavior is shaped by a pair of scissors whose two blades are the structure of task environments and the computational capabilities of the actor" (Simon 1990, p.7). My working hypothesis is that emotions relate to the latter blade: they amount to domain-specific information-processing systems that offer cue-based sources of inference that economize on the agent's limited information and computational capabilities. Emotions are assumed to serve functional roles that enable agents (a) to detect an important choice task, (b) to mobilize their cognitive resources to seek alternative courses of action that resolve the decision problem at hand and (c) to regulate mental activities by which selection of a favorable choice option (given her or his perceived aspirations and goal priorities). Furthermore, emotions shape three domains of human affairs in a robust manner. The first refers to uncertainties about the future and the agent's preferences, goal priorities and concerns vis-à-vis her or his current ones. The second concerns situations in which the individuals have difficulty in assessing probabilities and outcomes associated with risky prospects. The third involves interpersonal conflicts of interest that require some coordination of expectations and behavioral strategies (e.g. problem of self-control and long-term commitment). Note that these three types of choice task relate to intertemporal choice, decision under risk, and prosocial choice, respectively. Given that the validity of our argument depends on empirical confirmation, the current work scrutinizes whether some models of choice that incorporate psychological factors

4 Satz and Ferejohn (1994) and Hausman (1992, 2001 and 2003) provide an overview of the often debated rational economic choice explanation and its popularity among economists.

(cognitive and emotional elements included) yield more adequate explanatory accounts of actual behavior than approaches that fail to do so.

4. Plan of the book

This dissertation is organized in three parts. Part I offers the reader a historical survey of the rise and fall of economic analysis cleansed from explicit psychological assumptions and presuppositions. Part II addresses some methodological questions (among other philosophical issues) that guide the construction of accounts of choice behavior that take emotions (and other cognitions) seriously. It sketches a theoretical model that purports to offer a two-level explanation of decision-making behavior. Part III examines models of choice that are built in search of improved explanations (and predictions) of actual behavior within three important domains of economic analysis: choice over time, decision-making under risk, and prosocial (cooperative) choice. Furthermore it analyzes whether a two-level account of decision-making can trace the causal chain of emotional and higher order events that may produce manifest choice behavior (including those behavior patterns called anomalies).

Chapter 2 suggests that models cleansed from explicit psychological presuppositions arise in response to some 19th century economists' perception of two philosophical problems, regarded as obstacles to theoretical progress. One is the problem of measurability of hedonic utility and the other is the difficulty in developing a positive economic theory of (market exchange) behavior committed to the debatable descriptive and normative implications of psychological hedonism. The chapter goes on to explain the contemporary resurrection of behavioral models of choice by reference to two other problems: one relating to the prediction of anomalous behavior patterns, the other to explaining satisfactorily how actual choices (anomalies included) are produced in the real world.

Chapter 3 briefly addresses some conceptual and metatheoretical issues that may partly explain the ongoing efforts to reform economic models of choice. By conceptual clarification and using examples from the economic literature, it advances the argument that behavioral economists' plea for more realistic assumptions are consistent with the

vision that some unrealistic assumptions enable (rather than undermine) construction of an explanatory model/theory of choice.

Chapter 4 outlines a theoretical model that purports to explain decision-making behavior at the levels of mind and brain. It goes on to argue that a two-level account unveils the major roles emotion and higher order cognition may play in the production of manifest choice behavior, and therefore it allows for explanatory progress as enhanced causal articulation (and scope expansion).

Chapters 5 to 7 offer detailed studies of explanations of decision-making behavior within three domains: intertemporal choice, decision under risk, and prosocial choice. These chapters serve two main purposes. One is to confirm the hypothesis that certain behavioral models - by taking emotions seriously - can offer more explanatory and general accounts of actual behavior than conventional utility models. The other is to show that a theoretical model of boundedly rational choice with insights from neuroscience and evolutionary psychology squares well with (empirical) models that describe cognitive and affective processes that are (causally) significant for actual choice behavior. Chapter 8 wraps up the overall argument, briefly discusses some potential implications of a two-level economic explanation of choice that takes emotions seriously and is based on evidence from neuroscience, and finally addresses some questions comprising an agenda for further research.

CHAPTER 2
A HISTORY OF THE FALL AND REVIVAL OF PSYCHOLOGICAL ECONOMIC EXPLANATION

> The economist need not envelop his own science in the hazes of ethics, psychology and metaphysics.
>
> Fisher, *Mathematical Investigations in Theory of Value and Prices*, 1892

> Economics devoid of psychology is doomed to sterility (…). The methodological questions surrounding the crisis of the "hedonistic" model were never really resolved properly. (…) Economists who had learned that they were independent of psychology simply stopped thinking about the realism of micro-assumptions (after Friedman 1953, my emphasis).
>
> Shira Lewin, "Economics and Psychology", *Journal of Economic Literature*, 1996

1. Introduction

Economists among other behavioral scientists are often told that one merit of mainstream economic analysis is its capacity to offer parsimonious accounts of choice behavior that dispense with details of the complex machinery for decision-making. Despite the existence of many critics of the behavioral foundations of economic theory of choice (Katona 1954, Simon 1955, Sen 1990), there is a long-standing tradition of regarding the goodness of an economic model or theory as dependent on its predictive power (regardless of its underlying unrealistic assumptions about agents and silence about causal processes or mechanisms for behavior).[5]

All this seems to intrigue contemporary economists. It is necessary to find a solution to a methodological confusion concerning the independence of economics in relation to psychology and the possibility of developing satisfactory economic explanations of

5 See Friedman (1953) and Stigler (1986).

behavior cleansed from psychological presuppositions. This chapter attempts to be an exercise of rational reconstruction of ideas.[6] It aims to offer an interpretation of why many economists of the 20th century followed Fisher's ideas on the self-sufficiency of economic analysis and worked hard to develop accounts of behavior cleansed from psychology (at least explicit presuppositions, assumptions or doctrines). Based on a detailed literature about the shifting relations between economics and psychology, I also offer an account of the revival of economic psychological explanations after the 1980s (Coats 1976, Camerer 1995, Lewin 1996, Rabin 1998, Starmer 2000 and Camerer and Loewenstein 2004).

My overall argument, in this chapter, is built on the premise that a recurrent challenge posed to economists is to develop an account of choice behavior that meets its purported cognitive purposes. I advance the idea that two methodological problems partly explain why economists worked on economic models purged of (clear) psychology, whereas two other troubling methodological problems prompt contemporary decision researchers to argue for models of choice with more solid psychological foundations (than those of the expected utility maximization framework). I shall therefore look at the standard accounts of human action offered by late-19th century economists and raise two methodological difficulties: one is the "measurability of cardinal utility" and other concerns reliance on psychological hedonism as a basis for economic accounts of behavior. Together, I suggest, the foregoing problems seem to offer reasons for purging economic accounts from psychology. They are viewed as obstacles to the goal of transforming economics into an objective, rigorous (i.e. mathematically tractable) and predictively successful science just like physics. The current work takes up two other philosophical problems that partly account for the resurrection of psychological economic accounts: the difficulty in predicting behavior patterns that systematically and significantly deviate from predictions made by neoclassical choice theory and the difficulty in building an analytical model that yields scientific understanding of actual decision-making behavior. Inspired by the literature of philosophy of economics and economic methodology, the

6The method owes much to Professor Blaug's ideas on historiography (see Blaug 1997, introduction and 2001).

abovementioned problems are dubbed (a) prediction of anomalies and (b) identification of significant processes and mechanisms for behavior.

The remainder of this chapter is organized as follows. Section 2 discusses how two important late-19th century economists, Stanley Jevons (1871) and Francis Edgeworth (1881), relied on psychological hedonism to account for behavior within the realm of market exchange. Section 3 highlights certain methodological difficulties posed by Benthamite utilitarianism as well as by efforts to develop an alternative to the notion of cardinal (hedonic) utility that prompted economists to work on economic accounts cleansed from explicit psychology. Section 4 discusses the attempt to purge utility theory of psychological presuppositions, and in particular the work of some important early 20th century economists such as Vilfredo Pareto (1900, 1901), Eugene von Slutsky (1915), Frank Knight (1925) and Paul Samuelson (1938). Section 5 shows how economic theorists came to explain behavior without an appeal to intuitive psychology. Section 6 discusses why and how prediction of anomalies and identification of processes or mechanisms for behavior partly guide the resurrection of psychological economic explanations. Section 7 summarizes the main points addressed and draws three main conclusions about the fall and revival of 'economics with psychology'.

2. Economics and psychological hedonism

In his *Introduction to the Principles of Morals and Legislation*, Jeremy Bentham postulates that the pursuit of pleasure (and avoidance of pain) explains all human action and even serves as a criterion to evaluate the propriety of behavior. His ideas gave rise to accounts in terms of the hedonic utility principle. According to him,

> Nature has placed mankind under the governance of two masters, pain and pleasure… It is for them… alone… to point out what we ought to do, as well as what we shall do, in all we see, in all we think. (Bentham 1789, p.1)

Bentham's psychological hedonist framework ascribes an important role to utility in the explanation of any instance of human behavior. The very concept of utility was regarded as a cardinal measure of pleasure, happiness among other affective experiences. [7]

7Currently utility refers to a mathematical function that represents a rational agent's preference ordering.

Disutility in turn referred to negative affects that individuals avoided (e.g. pain). Note that, in such perspective, utility and disutility amount to properties of any object or commodity. The very notion of the rationality of human action comes to depend on whether behavior conforms to the principle of maximum utility (the net balance of the agent's pursuit of pleasure, benefits and avoidance of pain). Given the importance of this notion for leading 19th century economists, the appeal of Bentham's approach to economic action is worth examining.

2.1 Hedonic utility and explanation of economic facts

Unlike the classical tradition of Political Economy, which invoked labor value theory to address the issues of economic growth and wealth distribution, many economists in the 1870s took Bentham's doctrine of psychological hedonism as an alternative value theory that gives a rigorous and analytically tractable (i.e. mathematical) account of phenomena such as prices, resource allocation and market behavior. This shift in the *explanantia* and *explananda* was one of the hallmarks of the Marginal Revolution in economic thought.[8]

The first generation of marginalist economists used Bentham's notion of (cardinal) utility as an attempt to build a mathematically tractable account of behavior that would make economics as scientific as physics. Some 19th century economists saw the principle of hedonic calculus as a universal law of human nature from which economic relations and (market) exchange behaviors could be (scientifically) deduced. For instance, Jevons (1871/1970) drew on Bentham's hedonic utility to explain market phenomena in terms of pleasure-pain calculus, and hypothesized that the decreasing marginal utility principle was analogous to certain regular principles of physics (mechanics). In his own words,

> The nature of wealth and value is explained by the consideration of indefinitely small amounts of pleasure and pain, just as the theory of statics is made to rest upon the equality of indefinitely small amounts of energy (1871, Preface, p. viii).

8 Many economists have discussed the marginal revolution in economics. For details, see Blaug's, Coats' and Shackle's articles from the 1972 special issue of History of Political Economy reprinted in Black, Coats and Goodwin (1973). Blaug (1997) also provides an extensive analysis of the 1870s revolution in economic thought, methodology, and theorizing.

According to Jevons, value is determined by hedonic utility. The third chapter of his *Theory of Political Economy* suggests that pleasure and pain are the ultimate objects of economics, which is thus reduced to pleasure maximization. Jevons' account of behavior is in terms of hedonic calculus. To him, this psychological presupposition is necessary to make economics a scientific (and necessarily mathematical) endeavor.[9] The objective of every man is thought to be utility maximization in the direction of pleasure (1871/1970, p.97). Provided that utility is measured by increases in individual happiness or pleasure, Jevons' formal economic explanation becomes the formal counterpart of hedonic calculation. He claims:

> It is surely obvious that economics rests on the laws of human enjoyment, if those laws are developed by no other science, they must be developed by economists (1970, p.102)

Note that Jevons' analysis is committed to the doctrine of psychological hedonism, his line of argument premised on the idea that human behavior may be explained by the common denominator of utility calculus (maximization of pleasure, happiness and other positive affect).[10] However, Jevons acknowledges that hedonic calculus may not grasp all the complexity of human action and behavior. He suggests that the principle of (hedonic) utility maximization is a simplification of reality that serves the purposes of economics i.e. to explain market exchange behavior (Jevons 1871/1970, p.93).

Similarly, Francis Edgeworth assumes that utility amounts to happiness and pleasure as measurable magnitudes. In his article "Hedonical Calculus", Edgeworth appeals to the Benthamite thesis:

> Pleasure is measurable; and all pleasures are commensurable; so much of one sort of pleasure felt by one sentient being equateable to so much of other sorts of pleasure felt by other sentients (1879, p.395).

Three years later, Edgeworth's *Mathematical Psychics* applied quantitative methods – and, in particular, the conceptions and methods of physics - to Political Economy and other moral sciences for the first time so as to make economic analysis scientific. Taking

9 Jevons argued that economics had to be a mathematical science because it dealt with quantities.

10 Sugden (1991), p.754

such a step was welcomed by some of Jevons' contemporaries such as Jevons (1881) and Marshall (1881/1975). Edgeworth also postulates that the first principle of economics is that individuals are actuated solely by their self-interest. While admitting that behavior is not only determined by one's own interests and pursuit of pleasures, he argues that this "principle" is a proper simplification for the development of a theory that meets mainly the purposes of explanation of exchange behavior in a commercial society.

Note that Jevons and Edgeworth had much in common. Both used Bentham's notion of cardinal hedonic utility to construct an improved explanation of behavior. Edgeworth built on Jevons' approach to describe the mechanics of market exchange and under what conditions contracts occur. Both applied quantitative methods (e.g. differential calculus) to their studies to make economics an abstract, deductive and rigorous science like physics.

3. Motives for the 'de-psychologizing' of economic analysis

This section suggests that grounding economics on hedonic utility theory brought two methodological problems that constrain the economists' chance to develop a discipline with the scientific credentials of physics. They concern the measurability of cardinal utility and usage of psychological hedonism as basis for economic accounts of market phenomena. These two difficulties offer reasons for economists to develop economic accounts purged of psychology. I do not argue that measurability of utility and the using psychological hedonism as foundation for positive and normative economic analysis are the actual or sufficient causes of the fall of psychological economic explanations in the late 19th century. Instead, my point is to suggest that the abovementioned methodological problems partly do the explaining of why economic analysis centered on the hedonic calculus hypothesis lost its appeal and motivated building of a science of economics free from psychology.

3.1 Challenges for hedonic utility analysis

The first generation of marginalist economists remained quite enthusiastic about the prospect of theoretical progress associated with hedonic utility analysis.[11] In a sense, they did not properly investigate some problems associated with cardinal utility analysis drawing on psychological hedonism. One concerns the existence of an objective measure of agent's pleasures and pains (i.e. sources of individual utility and disutility). Another has to do with the use of psychological hedonism as a basis for positive as well as normative accounts of behavior regularities within the economy.

According to Stigler (1950), Jevons struggled to deal with the problem of utility measurability. In the first edition of his *Theory*, he denied the possibility of measuring utility:

> I have granted that we can hardly form the conception of a unit of pleasure or pain, so that the numerical expression of quantities of feelings seems to be out of question" (1871, p. 12)

Provided that the advancement of utility analysis depended on the existence and measurability of utility, Jevons and his contemporaries proposed that utility of a good to an individual could be estimated by using the amount of money that she or he spent to acquire the good.

> It is from the quantitative effects of the feelings that we must estimate their comparative amounts …when a man has purchased enough, he would derive equal pleasure from the possession of a small quantity more as he would from the money price of it (1879, p.11)

Note that an effective solution to this methodological problem would require an accurate way of measuring pleasure and pain. Economists of Jevons' time knew that nothing like one hedonimeter existed. In response to that they downplayed the significance of the measurability problem. Some marginalist economists went to argue that the issue of measurability of utility was not essential to meet their purpose of deriving demand curves (Stigler 1950b). Yet we can think of some applications to hedonic (cardinal) utility

11 For instance, Jevons applied his analytical account to show that parties involved in exchange gain satisfaction (i.e. improve their welfare). Menger, in his turn, stressed the generality and structural elegance of utility theory by explaining real-life phenomena, such as exchange behavior, wage and overall price determination (Stigler 1950, p.320).

approach that highlight the constraints on theorizing and explanation put by the issue of measurability. For instance, using hedonic utility to make welfare analysis would bring the problem of estimating the utility of a good to an individual and to compare it with affections (pleasures and pains) experienced by others. A second (related) methodological problem is reliance of psychological hedonism as a basis for economic positive and normative accounts.[12]

Grounding economic explanations on Benthamite utilitarianism had brought some embarrassment to economists. The first objections to the doctrine of psychological hedonism came from psychologists. Based on theoretical and empirical studies, they ridiculed the principle of pain-pleasure calculus and called the doctrine of psychological hedonism unscientific. In his *Principles of Psychology*, William James argued that a proper explanation of behavior could not be reduced to hedonic calculus, and found no empirical evidence that behavior was driven only by the human goal of maximizing pleasure and minimizing pain (pp. 551-5). Other psychologists (e.g. William McDougall 1910) went on to criticize economic analysis because it was built on a false psychological principle.

Following the methodological visions put forth by Neville Keynes, most 19th century economists did not take psychologist's critical remarks seriously. Even though they agreed that there were many reasons for action other than pleasure maximization, marginalist economists argued that omissions and simplifications were inevitable (Lewin 1996, p. 1300). Institutionalist economists in turn drew upon criticisms made by psychologists in order to suggest changes in standard economics. Thorstein Veblen (1909) and Ezekiel Downey (1910) argued that standard economic theory failed to explain important phenomena (i.e. the rise and persistent of institutions and economic change) due to its focus on pain-pleasure calculus.[13] All the above attacks to economic analysis gave rise to disputes within the profession. In response to such problems, I

12 For details, see Sen (1990) and Hollis and Sugden (1993).

13 A thorough analysis of the institutional economists' critique of marginalist economics grounded on hedonic psychology can be found in Lewin (1996).

claim, descents of marginalist revolution engage in building a science of economics cleansed of hedonic psychology.[14] Irving Fisher, for example, suggested that utility were independent of psychology and philosophical entities. To him, economists ought to content themselves with the postulate that each individual acts as he desires (1892, p.11). As he put it,

> This foisting of Psychology on Economics seems to be inappropriate and vicious … to fix the idea of utility the economist should go no further than is serviceable in explaining economic facts. It is not his province to build a theory of psychology" (1892, pp. 5, 11)

Note that the economists' task, from Fisher's point of view, would be to explain phenomena (e.g. exchange and price determination) that dispense with any specific psychological doctrine. In this case, economists should not waste their time working on psychological and philosophical matters.

In short, two methodological problems – finding a proper measure to hedonic utility and using hedonism as basis for positive economic analysis – can be interpreted as reasons for the developing a choice-based utility theory.

4. How economists freed their accounts from psychological presuppositions and doctrines

Purging economic explanation of a debatable psychological doctrine seemed to involve two stages. One was the gradual shift from theorizing on cardinal hedonic utility to ordinal utility; another was building an explanation of behavior in terms of utility and probability axioms free of any explicit psychological basis.

14 Wesley Mitchell (1916) suggests that even Alfred Marshall eliminated the hedonistic language of his account of utility and consumer behavior (p.144-145).

4.1 Ordinal utility theory as a first step towards an "economic analysis without psychology" [15]

In response to the philosophical problems underlying hedonic utility theory, economists at the turn of 20th century 'renounced' psychological hedonism. Instead, they worked hard on developing a choice-based utility approach. The latter is thought to offer an answer to the question of how to make utility objective and formally tractable just like force and energy in physics and a 'solution' to the problem of measurability of hedonic utility.

4.1.1 Pareto's contributions

Following Fisher, the economist Vilfredo Pareto (1900 a, 1900b and1906) contributed to the development of the modern account of utility as an index of individual preference ordering. He rejected the hedonist approach to behavior and argued that economics aimed to explain facts such as market exchange and price determination.

Pareto's *Cours* reflects his dissatisfaction with the vague notion of utility. He agreed on Walras' opinion that utility is not a measurable thing in practice. This interpretation led him to elaborate on an anti-hedonistic perspective. In a letter to Benedetto Croce, he wrote: "nobody is capable of measuring pleasure. Who can say what pleasure is double another pleasure?"(1900, p.183)

However, Pareto tried hard to give a mathematical treatment to utility theory. He sought to make economic analysis rigorous by devising a notion of utility purged of psychological presuppositions and doctrines. He proposed to measure the value of utility indirectly through observation of actual economic phenomena. This methodological procedure, he argued, resembled physicists' methods for determining the length of light waves through the observation of optical phenomena (see Marchionatti and Gambino 1997, p.1334). Pareto even went on to suggest replacing the term utility with ophelimity, but was somewhat reluctant to abandon the notion of utility**.**

15 Economists have tried to offer models free from psychological assumptions but there are doubts as to their success in this respect. For details, see Sugden (1991).

> Until now, in order to establish economic doctrines we went back to choice. Choice has been explained as man's aim to achieve maximum pleasure. Between two things, man chooses the one that provides more pleasure. (…) The use of this point of view forces us to consider pleasure as a quantity. And this is what economists have established pure economic theories have done, and what we ourselves have done in the *Cours*: but we must admit that this is not a thoroughly rigorous method (Pareto 1900b, p.221)

In order to provide a formally tractable economic explanation, Pareto provided a utility theory grounded in human experience (actual behavior). [16] By analyzing empirically derived indifference curves, he deduced a mathematical function that yields an index of the curves that represent an individual preference ordering. According to Pareto, this formal treatment of utility makes economics 'more scientific'. In his own words,

> This entire theory… rests only on a fact of experience, that is to say, on the determination of the quantities of goods which constitute combinations between which the individual is indifferent…the theory of economic science acquires the rigor of rational mechanics; it deduces its results from experience, without bringing in any metaphysical entity (1906, p.113).

Pareto's proposal of ordinal utility yields an account of behavior patterns without any appeal to subjective factors.[17] The theoretical representation sufficed for the purposes of economics, i.e., predicting market (exchange) phenomena.

Pareto's analysis also drew on the mathematician Poincaré, who saw 'preference' as crucial to an empirically grounded study of utility. In a sense, Pareto succeeded in devising an explanation that was not committed to hedonism. Alternatively he proposed a choice-based treatment. In his "Summary" (*Sunto*), he explains why a systematic study of economic facts can dispense with hedonic psychology (and philosophy too):

> Pure economic equations simply express the fact of a choice, and can be obtained independently of the notion of pleasure and pain…For us, it is sufficient to note the fact of individual choice, without investigating the psychological or metaphysical implications of such a choice…We do not inquire into the causes of man's actions: the observation of the fact itself is sufficient… Pure economic equations and their consequences exist unchanged whether we start from the consideration of pleasure as a quantity, or we limit our investigation… exclusively to the fact of choice (Pareto 1900b, p. 221-224).

So Pareto took the first step towards the development of utility theory purged of the problematic implications of hedonism. His ideas gave boost to formally tractable analysis of exchange behavior. With Pareto's contributions economists could advance their

16 See Hollis and Sugden (1993).

17 His fellow economists were advised to leave the task of studying the true nature of value to others (1901, p.204).

capacity to represent theoretically phenomena within the economy without incurring in the problems of measurability of pleasure and pain. Instead economists took the first step towards an explanation of market behavior in terms of observable variables (instead of inscrutable elements).

4.1.2 The development of ordinal utility framework

Eugene Slutsky (1915) followed Pareto and argued that economic explanations could not have a firm scientific basis unless it dispensed with psychological assumptions and metaphysical hypotheses (p.28). With that in mind, he argued for an account of utility that referred to an objective scale of individual preferences. Slutsky derived his utility theory from mathematical properties of indifference curves and his analysis in terms of ordinal utility facilitated the development of a satisfactory account of price determination and optimal resource allocation, as well as ridding economics of the embarrassment of hedonic psychology. In a historical study of utility theory, George Stigler (1950) argued that Slutsky had changed the direction of economic explanations to the extent that his analytical developments yielded a utility theory that does not assign any important role to introspection (p. 382-383).

Like other economists of the early 20th century Frank Knight acknowledges the problems of economic accounts that are committed to controversial psychological presuppositions and doctrines (e.g. psychological hedonism). In his insightful 1925 article, Knight advances the idea that the basic difficulty in economic theory is the philosophical problem of explanation and its relation to human behavior (p.372). To him, a choice-based approach to utility is necessary for a positive economic account. Knight even suggests that behaviorism is the only psychological thesis that squares well with economic theory. [18]

As a philosophical doctrine, behaviorism holds the thesis that explanation of behavior dispenses with any reference to mental entities and processes (Graham 2005, p. 2). In such perspective, a scientific explanation is only in terms of empirically observable

18 As Knight puts it, "the economist as a scientist may adhere to behaviorism (1925, p.388).

(external) factors rather than philosophical or psychological elements. [19] Behaviorism might fit with some economists' vision that freeing utility theory from psychological hedonism would allow for a positive and formally tractable account of market exchange phenomena. Perhaps this might explain (at least partly) why a behavioristic perspective became so popular among later generations of marginal economists.

Furthering Slutsky's and Knight's theoretical developments, Hicks and Allen's (1934) work advanced an ordinalist approach to behavior that further dispensed with introspection and psychologism by replacing the concept of marginal utility with that of marginal rate of substitution. Hicks and Allen attempted to devise an account in terms of the marginal rate of substitution principle. Economists welcomed their efforts to offer an explanation of behavior purged of a psychological foundation. However, some remained skeptical in relation to Hicks and Allen's analytical treatment being sufficient; their empirical studies rested on ambiguous results and therefore left room for psychological interpretations (Samuelson 1938, Lewin 1996).

4.1.3 Samuelson's contributions

Samuelson attempted to overcome Hicks and Allen's problem by "dropping off the last vestiges [from psychology] of the utility analysis" (1938, p.62, bracket is mine). Driven by his operationalist methodological convictions, he tried to rid utility analysis of the last psychological presuppositions by studying preference ordering and indifference curves based upon actually observed choices.[20] His account of revealed preference was deduced from a set of axioms of rational behavior in a choice setting in which agents were assumed to make internally consistent choices.

Samuelson hypothesized that if a bundle x were revealed preferred to another bundle y, there would never be cases in which y will be preferred to x. This theoretical formulation

19 For an overview, see Watson (1930), Skinner (1953 and 1987), Stich (1984) and Zuriff (1985). Economists' discussion of behaviorism can be found in Sudgen (1991), Lewin (1996) and Hands (2001).

20 For details of Samuelson's operationalism, see Hands (2001).

yielded a natural link between the demand function and individual preferences (observed choices). Empirical demand functions were thought to highlight individual preferences. Hence, Samuelson's theory of revealed preference yielded an economic explanation of behavior that does not rest upon any explicit psychology. [21] Individuals are expected to choose what they want. This is because 'what they want' is, by hypothesis, 'what they (actually) chose'.

As we can see, economists in the first half of the 20th century made intellectual efforts to develop a choice-based approach to utility. Unlike the first generation of marginalists, economists tried hard to remove explicit psychological presuppositions from utility theory. Their emphasis on choices in the explanation of economic behavior revealed some behavioristic leanings. The debatable notion of hedonic utility gradually evolved to become an ordinal utility that refers to an index of a rational agent's preference ranking.

Samuelson's theory of revealed preference was a decisive move toward an explanation of behavior without reference to psychological presuppositions. Behavior is explained in terms of choice that is equivalent (by hypothesis) to the outcome of a utility maximization problem. With the revealed preference approach, the 'de-psychologizing' trend within economics came very close to its peak.

5. The crowning glory of economic accounts free of (explicit) psychology

The 'crowning glory' of economic explanations purged of any explicit psychology was reached when decision theorists devises an analytical framework that could be applied to behavior under conditions of risk or uncertainty. This section discusses the move to an economic explanation of risky behavior free of psychology and notes that Savage's (1954) variant of expected utility theory provided an economic analysis that dispensed with psychological foundations, since his explanation of behavior was centered on quasi-logical relations (rational choice axioms and subjective probability rules).

21 Yet it is possible to argue that Samuelson's approach is compatible with a behavioristic perspective. See Lewin (1996).

5.1 The origins of expected utility theory

Risky choice is a complex phenomenon that has troubled social and behavioral scientists. The origins of expected utility theory go back to the 18th century, when Daniel Bernoulli (1738) developed a theory to explain the famous St. Petersburg paradox. Based on a gamble that yields an infinite monetary value, Bernoulli challenged the view that an individual was willing to pay to take part in a gamble in an amount up to its expected (monetary) payoff. [22] He argued that the agent's willingness to pay to gamble revealed the expected value or utility she or he ascribed to it (rather than its expected monetary value). On this basis, Bernoulli developed his account of choice under uncertainty. His theory hypothesizes that the value of a gamble corresponds to the weighted sum of expected utilities of its prospects (i.e. the sum of the multiplication of utilities of each outcome and the probabilities associated with them in every state of nature).[23] From this perspective, individuals act in conformity with the expected utility maximization principle.

5.2 von Neumann and Morgenstern's axiomatization of expected utility analysis

In the 1940s, economists became very interested in exploring Bernoulli's treatment of expected utility so as to better explain actual risky choice and its puzzles. Yet, they were still troubled by the fact that the expected utility approach required a notion of numerical (cardinal) utility that seemed to conflict with their developed notion of ordinal utility (Starmer 2000, p. 334).

The mathematical economists John von Neumann and Oskar Morgenstern (1944/1947) found an ingenious way of ensuring correspondence between the required cardinal scale and the ordinal utility function by formulating a set of axioms that allowed them to derive an expected utility theory. Like Bernoulli, von Neumann and Morgenstern assumed that inferences about expected utilities were drawn from the objective probability and utility attached to the prospects. Their expected utility theory differed from Bernoulli's approach in that it was deduced from axioms and relied on the hypothesis that individuals

22 The gamble is formulated as follows. A coin is flipped repeatedly until the first head is obtained. Those joining the lottery will have a payoff of $ 2n, where n is the number of coin tosses associated with the first head.

23 Bernoulli demonstrates that a rational person prefers gamble L1 to gamble L2, if and only if, the expected utility of L1 is as large as that of L2.

extract different utility measures from gambles (for details, see Starmer 2000). In this analytical approach, the derived expected utility function is unique to a positive linear transformation. Numerical utility values may be assigned to gamble payoffs that preserve the agent's preference ordering. Preferable alternatives are thought to yield high cardinal utilities (i.e. numerical indexes) vis-à-vis less favorable prospects. Put somewhat differently, vNM utility theory axioms give rise to a cardinal utility approach in which individual's preferences for risky prospects are ranked by their expected utility values.

The derived theoretical approach implies that a pattern of choice satisfying the axioms will correspond to the decision output that in turn maximizes an individual's expected utility function. Note that the account of behavior underlying vNM expected utility theory seems to appeals to consistency constraints on choice.[24] Such technical conditions facilitate the derivation of an explanation of choice that does not rest on any particular psychological presupposition or doctrine – individuals are assumed to make choices as if they drew optimal inferences about objective probabilities and outcomes. This assumption is later recognized as a problematic idealization, since it seems to constrain the theory's capacity to explain and to predict economically significant patterns of behavior, such as behavior patterns that reveal simultaneous gambling and insurance among other empirical violations of expected utility analysis.

One source of objection to the vNM expected utility approach is the premise that individuals make evaluations based on objective probabilities. This is regarded as a limitation of the theory. Actual agents exhibit choice behaviors that cannot be predicted or explained by considering the simplification that individuals know the distribution probability function of all events given by nature. In response to that, Leonard Savage (1954) proposed an alternative analytical treatment based on the notion of subjective probability, deduced from agent's preferences for risky gambles. This tricky maneuver gave rise to subjective expected utility, often regarded as the "most brilliant axiomatic theory of utility ever developed" (Fishburn 1970, p.191).

24 For details, see Sugden (1991) and Hollis and Sugden (1993).

5.3 Savage's subjective expected utility model

Savage's *Foundations of Statistics* provided a variant of expected utility theory built on Frank Ramsey's (1926/1931) theory of subjective probability. In his perspective, probabilities are described as 'subjective' in the sense that they are derived from the agent's preferences over certain gambles. Probabilities reveal individual expectations (beliefs) about the outcomes associated with the perceived risky prospects.[25] Savage's expected utility theory is deduced from a small set of axioms (see chapter 6, for details). His axiomatic constructs are compared with rules of logic that allow for a practical guide to rational choice. In this sense, his theory serves a normative purpose. He argues:

> Pursuing the analogy of logic, the main use I would make of [the axioms] is normative, to police my own decisions and, where possible, to make complicated decisions depend on simpler ones (1954, p.20).

Savage's expected utility theory is put as a normative rather than a descriptive theory. Its role is to inform agents about what choice behaviors meet the axioms of rationality and as such maximize their expected utility functions.[26] This version of expected utility theory became very popular among behavioral and social scientists and was applied to explain a wide range of phenomenon. It is regarded as a general framework that carries the paradigmatic vision of human rationality, i.e., expected utility maximization of outcomes.[27] Subjective expected utility theory provides an explanation of behavior purged of any psychological content.[28] In light of this account, an individual chooses among risky prospects as if she or he maximized her or his expected utility (function).[29] Therefore, choice behavior is explained in terms of consistency (rationality) restrictions on individual preferences.

So Savage's approach seems to be the culmination of theoretical efforts seeking to purge psychological presuppositions from economics. In the wake of this approach, economists

25In other words, subjective probabilities seem to be embedded in a 'belief-driven approach'. See Sugden (1991) and Starmer (2000).

26 Although the normative interpretation of the subjective expected utility theory is very popular, there are also descriptive usages of it. For details, see Schoemaker (1982), Goldstein and Hogarth (1997) and Starmer (2000).

27 Details are found in Hogarth and Reder (1986).

28 According to Savage (1954, p. 17), preferences are revealed by "decisions between acts and not by response to introspective questions".

29 Friedman and Savage (1952), pp. 473-474.

can study all instances of behavior even in an uncertain world by only referring to a compact set of choice and probability axioms. Therefore, many regard the subjective expected utility (henceforth: SEU) framework as "the crowning glory of choice theory".[30] Despite the popularity of Savage's theory it is worth considering two methodological issues that can be viewed as reasons for incorporating 'psychology' into the body of economic theory.

6. The resurrection of psychological economic accounts

Expected utility theory already met with some criticism in the 1950s (e.g. Vickrey 1945, Baumol 1950, Katona 1951, Edwards 1954 and Simon 1955) as empirical evidence weighed against the assumption that risky 'decisions are made rationally with a view of maximizing the mathematical expectation of a utility function" (Vickrey 1945, p.327)

This source of criticism was not taken very seriously at the time. Most economists seemed to share Friedman's methodological views that the strength of expected utility theory depended on its 'predictive power' rather than its 'descriptive accuracy'. According to Friedman and Savage, the expected utility hypothesis would be rejected only if its predictions were contradicted by observation (1952, pp. 463-469). Since the general expected utility approach successfully was thought to predict many phenomena, many economists did not fret over responding to the early criticisms of expected utility framework on empirical grounds. However, the growing evidence that conventional theory of choice failed to predict economically relevant instances of behavior prompted some economists to acknowledge two methodological problems. One concerns prediction of anomalies, i.e., empirical regularities (behavior patterns) that cannot be easily dealt by standard choice theory. Another refers to the task of identifying significant processes or mechanisms by which actual patterns of choice behavior come about in the real world (apparent puzzles and paradoxes included).

30 Kreps (1988), p. 120.

6.1 Early evidence against the predictive power of expected utility theory

In the 1950s, Maurice Allais challenged the empirical validity of expected utility theory. He argued that the standard approach to risky choice provided a wrong (theoretical) representation of individual risky preferences. His main objection related to the independence axiom, which states that if two lotteries have an identical branch of probability and payoff, their levels of payoff or probability will not affect agent's choice (see chapter 6). Allais (1953) discusses two situations in which the above condition does not hold. He offers evidence that individuals often prefer a lottery that yields \$5m with a chance of 10%, i.e., L_{11}= (\$5m, .10) to another that gives 11% chance of winning \$1m and nothing with a chance of 89%, L_{12}= (\$1m, .11). He also shows that the same agents tend to prefer a sure gain of \$1m denoted by L_{21}= (\$1m, 1) to a gamble that yields \$5m with 10% probability and \$1m with 89% chance, i.e. L_{22}= (\$5m, .10; \$1m; .89).

On the basis of the independence axiom, the individual's preferences for winning 5mi would remain unaffected by changes in common consequence. Yet Allais showed that adding a chance of 89% chance at \$1m prompted a preference shift. This puzzle (or anomaly) within the mainstream body of analysis is called the 'common consequence effect'. Allais explains this empirical violation of independence axiom by suggesting that individuals are quite sensitive to changes in probability mass (for details, see Camerer 1995 and Wu 1998).

Allais discovered yet another empirically grounded phenomenon that violated the independence axiom - the 'common-ratio effect'. Individuals often prefer a sure gain of \$3000 L_{11}= (\$3000, 1) to a lottery that yields \$4000 with a chance of 80%, L_{12}= (\$4000, 0.80; \$0, 0.20). Yet they tend to prefer a gamble that yields \$4000 with a probability of 20%, i.e., L_{21} = (\$4000, 0.20; \$0, 0.80) to another that gives \$3000 with a chance of 25%. , or L_{22}= (\$3000, 0.25; \$0, 0.75). Allais tried to explain this phenomenon by suggesting that, unlike expected utility theory assumes, individual preferences between pairs of prospects respond significantly to probability changes. When individuals face two pairs

of gambles with a constant ratio of winning chances, they rely on the probabilities of the acts to make choices (Camerer 1995, Kuilen and Wakker 2004).[31]
In the late 1970s psychologists Daniel Kahneman, Amos Tversky, Paul Slovic among others designed experiments to further investigate the implications of expected utility theory for the purposes of prediction and explanation. For instance, Slovic and Tversky (1974) attempted to empirically test Savage's axioms of rational choice. In their experiments, subjects were told that their previous choices over risky prospects revealed Allais and Ellsberg paradoxes. Individuals were then allowed to change their choices. Slovic and Tversky found that most participants did not shift their preferences even when they had the chance to do so.

This experimental literature identifying anomalies, paradoxes and effects highlight two important challenges for expected utility theory – one is prediction of economically significant behaviors regarded as anomalies, another is construction of an alternative model or theory that better explains how actual behavior happens and under what conditions the so-called choice anomalies may occur (and disappear). Nonetheless, most economists of that time were not really worried about dealing with these issues. Let me explain why that was so.

Thomas Kuhn's (1962) and Imre Lakatos' (1970) ideas may be useful to understand why economists hesitated to abandon expected utility theory in the face of anomalies. Both philosophers of science emphasize that it is always possible to disqualify a disconfirming piece of empirical evidence by suggesting that the problem does not lie with a model or

31 Ellsberg paradox is a third famous example of empirical studies that challenge expected utility theory. In the early 1960s Daniel Ellsberg designed an experiment in which subjects were asked to decide between betting that a red or black ball would be drawn from a 100-ball urn (with 50 red balls and 50 black balls) and betting that a red or black ball would be taken from a 100 ball-urn (with unknown distribution of colors). He found that actual people often prefer to bet on a red ball for the urn with known distribution of balls rather than a red ball from the urn with unknown distribution. Individuals also prefer to bet on a black ball from the known urn rather than a black ball from the unknown urn. On the basis of expected utility approach, preferring bets associated with the urn with known probability means that the subjective probabilities associated with drawing red and black ball in this urn are greater than subjective probabilities of drawing red and black in the second urn. Yet, the probability of drawing red or black is one in both urns. The paradox arises because the subjective probabilities associated with urn 1 are greater than urn 2 and at the same time we know that they have to be equal to 1. Ellsberg tentatively accounts for this puzzle by suggesting that actual people dislike ambiguity. For details, see Ellsberg (1961) and Camerer (1995).

theory under study but with poorly designed experiments. [32] This complex issue refers to the Duhem-Quine thesis. The latter asserts that negative evidence against a theory does not offer compelling grounds for refutation because no theory is ever tested in isolation (it also involves a bunch of auxiliary assumptions). [33]

Furthermore, most economists were not concerned with the development of alternative models of choice that would describe processes or mechanisms underlying actual behavior. They left this cognitive task to psychology. This is because the standard expected utility framework was thought to make accurate predictions of a wide range of economic phenomena.[34]

6.2 The revival of economic accounts with explicit psychological presuppositions

By the late 1980s economists and psychologists had already collected various choice anomalies. It was only at that time that these pieces of evidence against the predictive and explanatory powers of choice theory motivated some economists to investigate whether their standard models of choice were built on strong idealizations and omissions that constrained their cognitive purposes and therefore ought to be reformed. This issue begs the question of why this did not happen before. One possible explanation is that advances in game theory and experimental methods allowed economists to investigate the empirical validity of expected utility and to compare the predictive powers of standard theory with alternative analytical accounts offered by other decision researchers like prospect theory (Starmer 1999 and 2000). As a result, economists could find that incorporating psychological variables into their models of choice would be worthwhile if this helped them to resolve economically relevant anomalies and eventually predict novel facts (Kahneman, Tversky, Knetsch and Thaler 1986).[35]

32 For an overview of arguments against the significance and robustness of anomalies as puzzles within the body of economic theory, see Frey (1991) and Starmer (1999).

33 For details, see Hands (2001).

34 In their two famous papers on expected utility theory and its sources of criticism, Friedman and Savage argued that economic analysis would continue offering satisfactory predictions of empirical regularities within the economy, regardless of its silence about processes or mechanism for behavior.

35 See Hertwig and Ortmann (2001) for an interesting discussion of differences in experimental practices of economists and psychologists.

The early psychological economic models of the 1990s seem to be extensions of the (expected) utility framework. Practicing economists adopt an incremental strategy so as to come up with a model that yields improved predictions of actual choice behavior. Their analytical accounts do not purport to challenge the core hypothesis of neoclassical choice theory; economists only add psychological assumptions that are thought to improve their predictions of actual behavior.

For instance, Machina (1982) provided an important generalized expected utility model that relied on specific properties of expected utility functions and indifference curves to build a model to predict patterns seen as choice anomalies, such as the common ratio and the common consequence effects. Various extensions of the expected utility model provide an account of empirical deviations from expected utility theory (e.g. violations of the independence axiom) by resorting to particular mathematical devices, such as indifference curves that fan out (Starmer 2000). There are also expected utility models that take psychological variables rather explicitly. Two examples are the models of regret and disappointment developed by Robert Sugden, David Bell and Graham Loomes. I shall discuss this issue in some detail in chapter 6.

Prediction of economically important choice anomalies seems to be the most important driving force behind the revival of the interdisciplinary field of psychological economics.[36] One may wonder why behavioral economists opted for an incremental reformist strategy. This may be because the profession prefers progress in small steps and variants of the utility framework are not really perceived as a threat (Rubinstein 2005, p.2).

There seems to be another reason for putting psychology and economics back together. It has to do with the problem of identification of process or mechanism that offers genuine understanding of how actual decision-making behavior happens in the world. There is evidence that some behavioral economists and decision researchers tried to address this complex issue. For instance, Herbert Simon (1955) challenged the standard expected utility model and developed a theoretical model that purported to explain actual behavior

36 The terms 'psychological economics' and 'behavioral economics' may be used interchangeably.

by uncovering causally powerful mental (cognitive) processes. In chapter 4, I describe Simon's model, which has been quite influential, and its explanation of behavior in terms of a sequence of heuristics (e.g. search, stopping search and selection processes).

Kahneman and Tversky's prospect theory (1979) is another theoretical attempt to provide account of actual risky choice behavior in terms of a sequence of cognitive processes called 'framing', 'evaluation' and 'editing'.[37] Unlike expected utility framework, choices over a prospect are determined by a value S-shaped function that is defined in terms of gains and losses.[38] Chapter 6 will show that prospect theory among other behavioral models of risky choice contributed to clearer understanding of various puzzling phenomena.[39]

More recently, important behavioral economists like George Loewenstein and Colin Camerer have tried to develop models of choice that shift from an incremental strategy from a process-description approach (Camerer, Loewenstein, Prelec and Slovic 2005). This may be due to some economist's recognition that uncovering hidden processes or mechanism for behavior might better explain why (and under what conditions) behavior patterns that systematically deviate from the standard benchmark of rational choice happen. The novelty of these models is that they draw on insights from neuroscience and cognitive psychology so as to offer understanding of how emotional and higher order

37 The first stage consists of individuals' interpretations of the problem situation, in which they relied on certain mental shortcuts to draw inferences about contingencies, or alternative acts and their consequences. Like the framing stage, evaluation is also dependent on how a situation is presented to the agent. Individual conceptual frameworks, norms, habits and values may shape this second stage of decision-making. The final stage constitutes of processes that enable individuals to further assess prospects and to select one with the highest value (i.e., greater than outcomes for other possible options).

38 Prospect theory assumes that low probabilities are overestimated, whereas moderate and high probabilities are under weighted. The outcomes of a prospect are represented as positive and negative deviations from a baseline or reference point (a neutral outcome).

39 Expected utility theory cannot predict important anomalies such as loss aversion and the endowment effect. The former (loss aversion) points to the fact that actual people care more about losing something rather than gaining something else of the same monetary value. The endowment effect refers to situations in which the amount individuals are willing to pay for something they do not have is less than the sum of money for which they would sell something they own. For details, see Thaler (1980, 1985 and 1992).

cognitive processes give rise to behavior. In so doing, they give to economic analysis a chance of explanatory progress in terms of improved causal articulation.[40]

7. Concluding Remarks

This chapter attempted to offer an interpretation of the decline and subsequent revival of economic explanation based on explicit psychological assumptions, presuppositions or doctrines. By reconstructing ideas posed by economists and decision theorists that have influenced our vision of the economic approach to behavior (and its shortcomings), I hypothesize that some methodological problems prompted followers of the Marginal Revolution to abandon the doctrine of psychological hedonism and to elaborate on accounts of behavior purged of psychological presuppositions. One refers to the issue of measurability of hedonic utility; another concerns the use of hedonism as a basis for positive and normative economic analysis with a level of objectivity and rigor similar to physics.

The first conclusion drawn from this chapter was that Samuelson's theory of revealed preference circumvents the problem of grounding economic theory on a particular psychological doctrine. His approach offers an explanation of behavior in terms of observed choices that meet formal consistency requirements. Samuelson explicitly claims that observed choices suffice to explain economic behavior. In his perspective, observed choices reveal the rational agent's inner states (e.g. preferences and desires). If this is so, Samuelson's analysis can fit to a behavioristic interpretation. In this case, it can be suggested that his revealed preference theory did not really remove all vestige of psychology from economic analysis.

The second conclusion is that Savage's subjective expected utility theory is regarded as the "crowning glory of choice theory' due to its ability to account for various phenomena

40 Based on a metatheoretical framework, chapter 3 will address the complex issue of reformist theorizing and its link with explanatory progress. In chapter 4, I try to contribute to research that adopt a process-description strategy by tentatively suggesting a theoretical model of choice that explains actual decision-making and the significant roles emotions may play in it.

without any explicit reference to psychological intuitions. Behavior is explained (predicted) by means of a compact set of rational utility and probability axioms. SEU theory is the culmination of theoretical efforts to free economics from psychological presuppositions. Nevertheless, some psychological presuppositions remain in Savage's utility framework. For instance, the independence axiom can be read as an implicit statement about the agent's ability to compare choice prospects.

My third conclusion is that decision researchers' willingness to resolve the problems of (a) prediction of anomalies and (b) identification of significant processes and mechanisms for behavior explain (at least partly) the revival of psychological economic accounts. All this comes to challenge the widespread view that economic analysis does not need a behavioral foundation. Rather, such theoretical developments are in tune with the economic methodologist's vision that recurrent and robust anomalies serve as triggers of scientific innovations. Then, I end this chapter by paraphrasing the economist Paul Schoemaker - today's empirical puzzles to conventional wisdom offer the 'seeds' for tomorrow's improved explanations of a complex phenomenon liked decision-making behavior.[41]

41 In Schoemaker's own words, "it is likely that today's paradoxes and persistent EU violations hold the seed of future normative as well as descriptive theories of choice. After all it was a paradox… that gave birth to the current normative model" (1982, p. 556).

CHAPTER 3
METATHEORETICAL FOUNDATIONS

> It is plainly and patently bad social science that we don't care about how realistic our assumptions are.
>
> Matthew Rabin (2002) "A Perspective on Psychology and Economics, *EER*, p. p.672.

1. Introduction

There is a long-standing tradition that evaluates the worth of an economic model or theory by its predictive ability, regardless of the unrealisticness of its assumptions.[42] In his famous essay, Milton Friedman coined the dictum 'the better the economic theory, the more unrealistic its assumptions' (1953, p. 11). These methodological ideas became quite influential in the profession. More recently, a growing number of economists call Friedman's metatheoretical ideas into question. Some of them seem to appeal to Simon's vision of boundedly rational choice to suggest that the goodness of a model or theory is dependent on whether its underlying assumptions are acceptably unrealistic in the sense of correctly representing those explanantia elements that are necessary and sufficient to deal with a particular explanandum phenomenon. According to economists like Colin Camerer and George Loewenstein, the conviction that increasing the "realism of economic assumptions" will improve the quality of economic models and theories is at the core of behavioral economic analysis (2004, p. 3). In this chapter, I aim to present and briefly analyze some methodological issues that guide behavioral economists' efforts to develop improved explanations of decision-making behavior. To accomplish this task, I draw on Uskali Mäki's method of isolation (1992, 1994 and 2004). His metatheoretical approach takes theory building and theoretical innovations as responses to problem situations stemming from questions about the sufficiency and/or necessity of explanantia

[42] See Mäki (1994 and 1998a), for a careful distinction between the term realism (a philosophical doctrine) and the term realisticness (an attribute of theoretical statements like models, theories, assumptions).

items for derivation of a model or theory that satisfactorily deals with the chosen explanandum phenomenon.

The current work suggests that economic models, theories or approaches involve theoretical isolations, which are effected through idealizing and simplifying assumptions. Theoretical isolation occurs whenever a set of factors is isolated from the influence of other aspects of the world. In such perspective, models and theories are built through theoretical isolations that serve to pick out certain explaining elements in order to deal with a particular (explanandum) phenomenon. This leads to the vision that explanation and isolation are tightly connected.

Different models (theories) are thought to carve out the world differently by specifying an isolated system that accommodates two sets – one of explaining items (e.g. processes, events, states, causal powers) and another of explained items (e.g. events, behavior patterns and empirical regularities). These are the explanantia set and explananda set, respectively. From this perspective, models can be taken as representations that involve manipulations (through the choice and usage of particular assumptions, for instance) that serve to isolate major properties and causal relations of a complex phenomenon under study from the influence of anything else in the world.[43] The present chapter suggests that a new model/theory of choice is developed when economists realize that the inclusion of new explanantia elements might improve the model or theory's capacity to meet its purported cognitive goals (e.g. predicting, explaining or understanding a selected explanandum phenomenon).

To be more specific, this chapter analyzes the metatheoretical foundations underlying innovations in behavioral economic accounts of choices. It makes three main points. The first is that there are two theorizing strategies that might be useful to understand the development of models in direction of improved predictions and explanations of the

[43] If this is so, it makes sense to suggest that some models resemble experiments. Both amount to representations of things in the real world by isolating elements that are regarded as playing major roles in production of the phenomenon under study from influence the rest of the world. For details, see Mäki (2005).

complex phenomenon of human decision-making. One is the 'de-isolation strategy'. This is a matter of increasing a theory or model's degree of approximation to reality by way of relaxing some unrealistic psychological assumptions and incorporating additional explanatory factors into the body of analysis (Mäki 2004, p. 339). Another is the re-isolation strategy. It differs from 'de-isolation' in the sense of replacing (rather than supplementing) one or more previously included explanatory variables by including one or more previously excluded explanantia factors. The second claim is that most behavioral economic models rely on the de-isolation strategy (rather than that on re-isolation). This is at least partly so because the re-isolation strategy carries some revolutionary implications; it proposes to replace theoretical statements playing major roles in the neoclassical account of behavior (e.g. rational choice behavior corresponds to the expected utility maximization principle) with empirically grounded statements (e.g. rational choice behavior is shaped by agent's computational powers and the informational structure of the task environment). Inspired by Rubinstein's (2005) interpretation of contemporary behavioral economic analysis, this chapter suggests that theoretical innovations in this field, most of the time, follow the strategy of de-isolation because the profession prefers achieving progress by small steps. The third point is that the ways in which behavioral economists employ the de-isolation and the re-isolation strategy reveal a realist goal of truthfully representing decision-making by its major constituents.

2. Methodology in neoclassical and behavioral economics

Some economists regard contemporary behavioral economic analysis as a natural continuation of neoclassical economics rather than a radical shift from it (Rabin 2002, Camerer and Loewenstein 2003 and Rubinstein 2005). However, some behavioral researchers criticize the expected utility maximization framework and even propose alternative accounts of choice behavior like prospect theory (Kahneman and Tversky 1979, Thaler 1992, Simon 1997). In this section, I advance the argument that various objections to neoclassical economic analysis have methodological roots. By comparing the methodology of neoclassical economics with that of behavioral economics, I show that behavioral models also deploy analytical methods to deal with behavior patterns. Yet they differ from standard expected utility models of choice in exposing a metatheoretical

perspective that squares better with Simon's realist orientations than with Friedman's instrumentalist leanings (Friedman and Savage 1948 and 1952).

2.1 The methodological basis of neoclassical economic analysis

Despite some disagreements on what mainstream economics is like, economists of the 20th century agree that the neoclassical account choice of behavior is in terms of the expected utility maximization hypothesis (Friedman and Savage 1948; see Sugden 1991 for a critical analysis). They acknowledge that the subjective expected utility framework is built on some psychologically unrealistic assumptions. The expected utility framework assumes that an individual makes a choice as if she or he picked out the course of action that maximizes the expected value of her or his utility function. Yet real-world agents do not have unconstrained computational facilities to draw optimal inferences about all available prospects and outcomes associated with them in every state of nature.

In his *Foundations of Statistics*, Leonard Savage emphasizes that subjective expected utility theory cannot be criticized for its underlying unrealistic assumptions because it does not purport to offer a description of actual human choices. Rather, it is a normative theory of choice behavior that informs agents about the course of action that yields rational (optimal) outcomes.[44] In their famous 1948 and 1952 papers, Friedman and Savage argue that the validity of subjective expected utility theory does not lie on the descriptive validity of its assumptions (Friedman and Savage 1948 and 1952). They argue:

> The validity of this assertion [expected utility maximization hypothesis, my brackets] does not depend on whether individuals know the precise odds, much less on what they say that they calculate and compare expected utilities or think that they do, or whether psychologists can uncover any evidence that they do, but solely that it yields …accurate predictions. (1948, p. 298)

As the passage indicates, what really matters is whether the theory or hypothesis offers accurate predictions, not whether theory draws on empirically supported premises. This is largely so because meaningful prediction is viewed as the ultimate goal of any positive

[44] Although many agree that (subjective) expected utility does not serve descriptive purposes, there are also some descriptive usages in the literature. The huge literature on decision research sometimes offers descriptive readings of expected utility theory. For details, see Schoemaker (1982) and Starmer (2000).

economic theory/model (Friedman 1953, p. 7). This emphasis on predictive accuracy as the criterion by which economists ought to evaluate the goodness of any economic model/theory allowed for an instrumentalist (anti-realist) interpretation of Friedman's theorizing. [45, 46] From this perspective, a predicted outcome is deduced from a compact set of utility axioms and probability rules. Explanation is supposed to have a similar structure - an explanandum phenomenon (phenomenon under study) is derived from initial conditions and assumptions about agents' preferences and expectations. In this case, a scientific explanation can be equated with an exercise of formal derivation of behavior patterns from a set of analytical statements. As a result the neoclassical economic explanation is regarded as prediction, except that the phenomenon under study was already observed. [47, 48] This allows for a quite minimalist view of explanation, which has very little to do with the intuitive vision of explanation as tracing the causes of an event or phenomenon under study (for details see Hausman 1992 and 2001).

2.2 Methodological foundations of behavioral analysis

Behavioral economic accounts emerged to challenge the expected utility framework. Some economists like William Baumol (1951), George Katona (1954) and Herbert Simon (1955) made serious objections to the way neoclassical economists theorized choice behavior. Their main source of criticism was based on the idea that various instances of human action contradicted the hypothesis of expected utility maximization. Against Friedman's methodological view - that the adequacy of a theory does not depend on the 'realism of assumptions'- early psychological economists proposed to develop models/theories built on empirically valid assumptions about actors.

[45] Yet there are interesting arguments that challenge the standard interpretation of Friedman as an anti-realist or instrumentalist about scientific models and theories. For details, see Mäki (1986, 1992 and 2003).
[46] Friedman and Savage are quite optimistic about the predictive abilities of expected utility theory. They even suggest that whatever the behavioral mechanisms whereby actual individuals make choices, these outcomes are consistent with the expected utility principle (Friedman and Savage 1948, p.298)
[47] See Hausman (2001) and Hands (2001) for a critical analysis of the conflation between explanation and prediction, also known as the symmetry thesis.
[48] Perhaps Marshall had similar ideas when he suggested that explanation in neoclassical economics resembles "prediction backwards".

Simon, for example, rejected Friedman's aphorism - the better the economic theory, the more unrealistic its assumptions - and complained that the methodology of neoclassical economists was based on Friedman's 'principle of unreality' that presupposed an asymmetrical relationship between assumptions and theory, which allows for mathematical derivation of an empirically valid theory from unrealistic assumptions (Archibald, Simon and Samuelson 1963, p.229). His main objection to the methodology of neoclassical economic accounts was that it took unrealistic assumptions as virtues in theorizing. Simon argued for a methodology based upon a distinctive theorizing principle called 'continuity of approximation' (ibid, p. 230).[49] He suggested that models of bounded rationality that uncover specialized heuristics and cognitive processes would yield genuine understanding of how problem solving and decision-making occur in the world (e.g. Simon 1991 and 1997).

Following Simon's methodological advice behavioral economists pursue the development of an economic model/theory that aims at a satisfactory approximation to reality given the purposes of prediction and/or explanation. Behavioral economists share the idea that certain unrealistic assumptions may constrain a model/theory's capacity to accomplish its tasks and therefore ought to be revised (e.g. Simon 1997, Camerer and Loewenstein 2003). They differ from the methodology of neoclassical analysis since behavioral researchers explicitly engage in investigating the predictive and explanatory consequences of their model's unrealistic assumptions (i.e. omissions and idealizations). When there is systematic evidence that the standard body of analysis cannot deal with important behavioral patterns, empirically grounded assumptions are incorporated in order to improve the model's predictive and/or explanatory capabilities.

It is important to stress that meaningful prediction is also a goal of behavioral economic theorizing. According to some behavioral researchers, it is worthwhile to add more explanatory factors to the standard model of choice (i.e. it does pay off to engage in a de-isolative strategy) if the reformed model can accommodate instances of choice behavior

[49] Unfortunately, Simon offers little detail about his alternative metatheory. For critical remarks of Simon's objections to idealizations by mainstream economics, see Nelson (2001).

already covered by conventional models as well as empirical regularities regarded as anomalies (Kahneman, Tversky, Knetsch and Thaler 1986). In this sense, behavioral economic analysis does not differ from neoclassical economics (influenced by Friedman's methodology). Nevertheless, there seem to be quite significant differences on theorizing and explanation. Behavioral researchers do not entertain the vision that prediction is the only goal of economic models/theories or that explanation is just like prediction of already known events. Scientific explanation is indeed viewed as redescription of elements (process, state, course of affairs) that are (causally) relevant to the occurrence of a phenomenon under study. In other words, their view of explanation seems to involve the task of unearthing processes or mechanisms whereby a chosen *explanandum* comes about in a complex world of dependency relations (Mäki 2003). Models that succeed in describing relevant processes and states underlying manifest behavior are thought to offer some understanding of how the explanandum relates to the causal structure of the world. Some behavioral economists drawing on a process-description approach seem to suggest that models unveiling hidden processes or mechanisms for behavior can contribute to predictive advancements in economic analysis, since meaningful predictions require some knowledge about the driving forces behind patterns of actual choice behavior (Camerer, Loewenstein and Prelec 2005).[50]

Another important source of difference between neoclassical and behavioral analysis lies on their visions regarding scientific theory and explanation. Unlike some neoclassical economists who seem to entertain an instrumentalist view of theorizing, behavioral researchers try hard to come up with models of choice that are congruent with reality. Even though their models of choice still rely on some idealizations and simplifications, behavioral economists employ the strategies of de-isolation and re-isolation in a way that is consistent with a realist posture on theorizing and explanation. This statement amounts to an empirical question. In the next four chapters, I show that proposals of de-isolation and re-isolation by behavioral economists expose their concern with the development of models/theories of choice that represent the complex reality of human thinking and decision-making by its major constituents. Hence, we may have grounds for suggesting

[50] For an interesting argument about this issue, see Hausman (2001).

that disagreements between neoclassical and behavioral economists derive from different methodological foundations.

3. Theorizing strategies and the search for improved explanations

This section further explores the incremental approach and the process-description strategy already presented in chapter 2. In particular, I draw attention to the importance of isolations in behavioral economic theorizing. In addition, I show that the incremental approach and the process-description approach can be regarded as efforts of de-isolation and re-isolation respectively within the explanantia set of the neoclassical model/theory of choice.

Recall that any model/theory involves isolation among the explaining and explained elements. For instance, a particular model/theory of choice isolates a potential explanantia set from the rest of possible explaining items. To simplify matters, consider that a chosen set of explaining items $\{s_1, \dots, s_k\}$ constitute the explanantia set S_{k-} (where an explaining item s_k is also included). This set of explaining items (S_{k-}) is isolated from all other possible explaining items denoted by $\{s_{k+1}, \dots\}$. The latter are thought to constitute another explanantia set, denoted by S_{k+}. Quite similarly, the items constituting the explanandum set Mn- are also isolated from other existing explanandum phenomena. These excluded items $\{m_{n+1}, \dots\}$ constitute a new set of explananda denoted by M_{n+}.[51] This perspective suggests that the very choice of the explanantia set S_k restricts the set of explananda phenomena M_{n-} containing $\{m_1, \dots m_n\}$ items. If this is so, a model or theory's capacity to meet the purposes of prediction and/or explanation is dependent on the theorist's choice over explaining factors.

Following chapter 2, I advance the argument that economists' detection of significant and recurrent choice anomalies prompt them to revise the previously isolated explanantia set characterizing the standard economic model of choice. Empirical anomalies motivate theorists to raise four types of questions (see Mäki 2004). The first question is whether the initially isolated items of the explanantia set (S_{k-}) are sufficient to deal with the

[51] For details, see Mäki (2004), pp.322-324.

researcher's chosen set of explanandum phenomena (M_{n-}). The second question is whether the items of the previously chosen explanantia set (S_{k-}) suffice to explain new explananda items from the set (M_{n+}). The third question concerns whether the items from the explanantia set (S_{k-}) are necessary to account for phenomena from the initially chosen explanandum set (M_{n-}). Finally, the fourth question is whether the explaining elements from the isolated explanantia set (S_{k-}) are necessary to explain other explananda phenomena from the M_{n+} set. Depending on the type of question that the behavioral economist tries to answer, two different types of reformist strategies will be adopted.

3.1 The incremental approach as a matter of de-isolation

Contemporary behavioral economists do not really deny that the (expected) utility framework is applicable to a wide range of phenomena (Rabin 2002 and Camerer and Loewenstein 2003). Rather, they seem to question the sufficiency of isolated explaining items to deal with instances of choice behavior that are regarded as anomalies (that remain unexplained) as well as new explananda items. This prompts the development of behavioral models that add explaining items (e.g. behavioral assumptions about agent's future discounting and attitudes towards radical uncertainty) to the utility approach without questioning previously included items of explanantia set (i.e. assumptions about individual's deliberate and careful judgment and decision-making and his or her rational preferences). This sort of theory building characterizes the incremental approach, and it can be interpreted as a matter of de-isolation.[52] In chapters 5 to 7, I show that the de-isolation strategy plays a prominent role in theorizing on intertemporal choice, decision under risk and prosocial choice. To illustrate, take the standard model of risky choice that is centered on the (subjective) expected utility maximization hypothesis. In response to expected utility anomalies (e.g. simultaneous gambling and insurance, certainty effect, reflection effect and overweighting of small probabilities), some behavioral researchers like Robert Sugden and Graham Loomes propose a variant of the expected utility model that adds 'regret' to the set of explaining items without questioning other items of the

[52] Camerer and Loewenstein explain how an incremental strategy of theorizing is put into motion:

> Modify one or two assumptions in the direction of greater "psychological realism". Often these departures are not radical at all because they relax simplifying assumptions that are not central to the economic approach (2004, p.3, my emphasis).

conventional body of economic analysis. In their model, regret amounts to individual's capacity to anticipate feelings of regret and rejoicing is an important factor affecting risky inferences and choices (Loomes and Sugden 1982, p.822). One positive implication of regret theory is that it can accommodate the abovementioned anomalies and therefore contributes to a reduced gap between actual and predicted behaviors.[53]

3.2 The process-description approach as a matter of re-isolation

Behavioral models of choice following an incremental approach brought explanatory and predictive improvements to behavioral economic analysis. Yet, the task of yielding understanding of how actual patterns of behavior come about and that of explaining the conditions under which recurrent anomalies happen gave rise to another theorizing strategy. It comes out of an answer to the question whether previously included explaining items S_{k-} is necessary to explain M_{n-} (i.e. those isolated items that the model/theory purports to deal with) and to cover M_{n+} (new explananda items). Provided that the answer to that question prompts the replacement of previously isolated explaining items with new explanantia ones, it gives room for the task of re-isolation (Mäki 2004).

Behavioral models following the above strategy challenges more clearly the standard body of analysis than those following an incremental approach. They substitute some explanatory factors (e.g. cognitive processes) playing a causal role in production of actual choice behavior for idealizing assumptions about agent's rational preferences and expectations (e.g. continuity axiom, sure thing principle). One distinctive trait of this theorizing approach concerns the emphasis put on description of (causally relevant) processes or mechanisms underlying patterns of actual choice behavior (anomalies included). Therefore, this theorizing approach is called process-description strategy.[54] Daniel Kahneman and Amos Tversky's prospect theory is an example of model of risky choice following a process-description strategy (Kahneman and Tversky, 1979). Prospect

[53] For details, see chapter 7 of the current work.

[54] Behavioral models of choice that follow a process-description strategy do not square very well with the standard expected utility framework. They often replace some isolated explanantia items of the conventional approach (e.g. utility axioms and norms of subjective probability) by incorporating new explaining elements (e.g. decision heuristics and domain-specific cognitive processes) so as to offer analytical accounts with superior degree of congruence with reality.

theory replaces the idealizing assumption of expected utility maximization by including explanantia items that are thought to play a significant role for the occurrence of actual choices under risk. According to this re-isolated account of risky choice, the agent is thought to estimate outcomes of available risky prospects on the basis of her or his perception of gains or losses, i.e., positive or deviations from a reference point (a neutral outcome). Prospect theory purports to represent a risky choice task in terms of an S-shaped value function in the domain of gains and losses (instead of a well-behaved expected utility function). This re-isolated account of decision under risk can explain and predict those behavior patterns that are already covered by the standard expected utility models as well as those regarded as choice anomalies (e.g. loss aversion and the endowment effect).[55] For that reason, prospect theory and its variants became quite appealing to decision theorists (e.g. Thaler 1981, Camerer, Rabin and Loewenstein 2003).[56]

3.3 Implications

The incremental approach and the process-description strategy attest the central role of isolations in building economic models and theories of choice. They also reveal behavioral economists' ongoing attempts to develop analytical models with greater predictive (and explanatory) abilities than standard accounts of choice. One distinctive trait of the reformist project characterizing behavioral economic analysis is the systematic study of whether certain models of choice, built on unrealistic assumptions, have isolated too much, too little or wrongly (given their cognitive purposes).[57] Unlike those endorsing Friedman's metatheoretical vision that the goodness of a model/theory does not depend on the empirical validity and truth value of its underlying assumptions, many behavioral

[55] Loss aversion highlights the fact that actual people care more about losing something rather than gaining something else with the same monetary value. The endowment effect refers to situations in which the amount an individual is willing to pay for something that she or he does not have is less than the price at which she or he would be disposed to sell something that she or he owns. See chapter 6 for more details of expected utility anomalies.

[56] More recently, some economists draw on insights from neurosciences in an attempt to develop behavioral models of choice that describe those emotional and higher order cognitive processes causally significant for decision-making behavior. In so doing, they seek to extract the explanatory gains of a process-description theorizing strategy.

[57] See Mäki (1994), for details on connections between isolation and explanation.

economists plead for models of choice with assumptions that are as realistic as possible (Simon 1997, Rabin 2002, Camerer, Loewenstein and Rabin 2004).[58]

Based on the method of isolation, it can be suggested that all economic models are necessarily unrealistic in the sense of being partial and incomplete (Simon 1997, Rabin 2002).[59] Assumptions are vehicles that serve to remove important features, properties and capacities of things in the world from the influence of anything else. With this in mind, it seems obvious that the behavioral economist's plea for models with realistic assumptions is not a defense of models cleansed from any simplification or idealizing assumption. Rather, it seems to be a call for systematic construction of empirically grounded models that offer adequate (theoretical) representations of the key aspects of actual decision-making within economically relevant domains of human affairs. Behavioral economists seem to employ reformist theorizing strategies of de-isolation and re-isolation since they are concerned about whether their models of choice have correctly isolated those factors that play major roles in the production of manifest behavior. This suggests that their reformist project is consistent with a realist posture towards theorizing and explanation.[60] The application of the method of isolation to the study of behavioral theorizing provides a clearer grasp of why models built on particular idealizing assumptions do not necessarily thwart the realist aim of science. Some unrealistic assumptions are compatible with the development of an approximately true account of choice behavior if they serve the purpose of isolating those (causally) significant factors about the phenomenon under study from the involvement of other elements that are thought to play negligible or

[58] In a recent paper, Matthew Rabin argues:

> Ceteris paribus, the more realistic our assumptions about economic actors, the better our economics. Hence, economists should aspire to make our assumptions about humans as psychologically realistic as possible (2002, p 658).

[59] Then, analytical models might violate 'the whole truth' and 'the nothing but the truth' principles. For details, see Mäki (2004).

[60] I am indebted to Caterina Marchionni and Ramón Fernandez for drawing my attention to the neutrality of de-isolation and re-isolation with respect to realism, instrumentalism and empiricism. For example, an instrumentalist about scientific theories can introduce a new explanatory item in order to improve the predictive accuracy of her model, which is in fact interpreted as being false or as having no truth value. Then, it would be a mistake to presuppose that any move in the explanantia set, via de-isolation or re-isolation, is done with the purpose of better approximating the mode/theory to reality.

unimportant roles.[61, 62, 63] Of course this issue also amounts to an empirical question. An effective way of knowing whether it is worthwhile to incorporate a new explaining item into the body of economic analysis via de-isolation or re-isolation is by comparing the effect that it exerts on the gap between predicted and actual behavior and/or on the model/theory's capacity to offer understanding of the (causal) chain of events that produce the phenomenon under study.

A third implication of the suggested metatheoretical framework is that it helps understand why re-isolative models are worth constructing and what their potential (explanatory) gains are like. Further development of re-isolative models (e.g. model of intertemporal choice that describes a projection bias and an account of risky choice in terms of affect heuristic) may carry some revolutionary implications to standard economic analysis. By replacing those explanatory items (that are thought to constrain the model's degree of congruence with reality) with new set of explaining elements (e.g. assumption that individuals are boundedly rational agents that make inferences and choices based on affect-driven heuristics), those behavioral accounts allow for improved descriptions of mental processes that might track the causal chain of events that produce manifest choice behavior (anomalies included). Therefore they contribute to progress as enhanced causal articulation and give economists the chance to address explicitly the complex issue of causation. [64]

[61] For instance, the idealizing assumption of positive time preference does not undermine development of an explanatory account of intertemporal choice. This theoretical statement can be paraphrased as a (correct or true) statement about the negligibility of an individual's structure of time preferences (it includes farsightedness, i.e., preferences discounted at negative rates) when intertemporal choice at the market level is considered. This assumption serves to simplify the complex structure of intertemporal preferences and contributes to (rather than undermines) derivation of an analytical model that approximates to the truth about intertemporal decision-making and its major constituents.

[62] As Mäki (2003) put the issue, "the violation of the whole truth does not imply the violation of nothing but the truth about some part of the whole" (p.499).

[63] For details on the different roles unrealistic assumptions play in theory building, see Musgrave (1981) and Mäki (2000).

[64] For further details on the complex issues of causation and explanation, see Salmon (1994). Hausman (2001) offers an interesting account of explanation in economics that guides the argument developed in this work.

4. Concluding remarks

This short chapter is an attempt to expose the metatheoretical foundations on which theoretical innovations in behavioral economic analysis are built. Based on some ideas on the cognitive goals pursued by practicing economists and insights from Mäki's framework, I provided a comparison between the methodology of neoclassical economics and that of behavioral economics (section 2). The first important lesson drawn from this chapter is that behavioral theorizing challenges Friedman's vision that economic models/ theories only aim at meaningful predictions (rather than descriptions of actual behavior) and therefore need not be built on realistic assumptions. Inspired by Simon's methodological ideas, behavioral researchers call the instrumentalist perspective on theorizing into question. Rather they propose ongoing empirical tests of whether some unrealistic assumptions allow for the derivation of a model/theory of choice that meets the purposes of prediction and/or explanation. If this is so, there are grounds for suggesting that behavioral economists pursue the development of accounts that theoretically represent the complex phenomenon of decision-making behavior more correctly (truly) than the standard expected utility model.

This chapter went on to present and briefly discuss two theorizing strategies that characterize contemporary behavioral accounts, the so-called incremental approach and the process-description approach (section 3). In light of the suggested metatheoretical framework, I argued that the incremental approach is a matter of de-isolation; it proposes to improve the quality of economic predictions by developing models of choice that supplement previously isolated explaining items of the mainstream body of theory with new explanantia items. Such de-isolative strategy is thought to motivate construction of some models that did a good job at accounting for some economically relevant choice anomalies (for details, see chapters 5 to 7). The process-description approach is in turn a matter of re-isolation. It deviates from the expected utility framework to the extent that its models/theories of choice replace one or more items of the selected explanantia with new explaining elements (e.g. cognitive and emotional processes or mechanisms).

The second lesson of the current chapter suggests that the development of models following an incremental strategy and a process-description approach highlight behavioral researchers' theoretical efforts to represent decision-making behavior by referring to its major constituents. This allows for a characterization of behavioral economic analysis as a reformist project motivated by doubts whether the resemblance between standard models of choice (e.g. discounted utility model, expected utility model and self-interest model) and reality are close enough given the purposes of explanation and prediction of actual behavior. Behavioral models are representations that attempt to isolate key factors of certain slices of the complex reality of actual decision-making. The current work emphasized that the behavioral economist's plea for economic accounts with more realistic psychological assumptions (than those underlying the expected utility framework) does not imply a naïve defense of analytical models of choice purged of idealizing and simplifying assumptions. Rather, the behavioral economist's discussion of certain unrealistic behavioral assumption serves to draw economist's attention to the need of empirical investigation of whether his or her theoretical statements about the determinants of human judgment and decision-making have not isolated too much, or too little, or do so wrongly. In so doing, behavioral accounts drawing on de-isolative or re-isolative strategies contribute to the development of analytical accounts that offer an improved grasp of the place of decision-making and its major constituents in the causal structure of the real economic world.

The final lesson of this chapter concerns the potential explanatory gains of building up models that draw on a process-description approach. Behavioral models that uncover those processes with major roles in the production of actual choice behavior (anomalies included) better deal with a complex problem posed to economics, the issue of causation (Simon 1997, Hausman 2001). Models of choice that describe processes might identify the (causal) chain of mental events that give rise to behavior patterns. In so doing, they yield a genuine causal explanation of decision-making that informs us economists among other behavioral researches about how emotion, higher order cognition and manifest choice behavior are related in a complex world of causal interdependencies. One important implication of a reformed model of choice that allows for an 'causal

explanation by mechanism' is that it may help decision researchers to distinguish significant and robust processes for behavior from some identified spurious correlations and eventually make accurate predictions and diagnoses of new phenomena (Salmon 1984, Glennan 1996, Elster 1999). Until now we do not have such a theoretical model that identifies how elements of the explanantia set work together to account for the explanandum phenomenon. This may put a limit on explanatory progress in economic analysis and may even partly explain why economists hesitate to talk about explanation as a cognitive task that cites causes (Simon 1997). I take behavioral economists to endorse the principle that an economic explanation in this richer, causal sense is worth developing because it provides that required knowledge for effective predictions of behavior and also informs us about how to behave to bring about certain choice patterns or to prevent them (Hausman 2001, p. 119).

In the next chapter, I attempt to further explore the above ideas by suggesting a theoretical model that purports to yield a mechanistic explanation of decision-making behavior at two complementary levels of analysis, mental and neural.

CHAPTER 4
A TWO-LEVEL EXPLANATION OF DECISION-MAKING BEHAVIOR

> In the behavioral model, by contrast, the choice of problems for the agenda is a matter of central importance, and emotion may play a large role in that choice… A behavioral theory of rationality (…) does not dissociate emotion from human thought, nor does it in any respect underestimate the powerful effects of emotion in setting the agenda for human problem solving.
>
> Simon, *Reason in Human Affairs,* 1983

> It is essential to explain complex phenomena at several levels, symbolic as well as physiological; complementary, not competitive.
>
> Simon, "What is an Explanation of Behavior?" *Psychological Science*, 1992

1. Introduction

Decision-making is a complex phenomenon in need of a clearer understanding than that provided by the basic (expected) utility framework. In the previous chapter, I argued that reformist theorizing strategies pursued by behavioral researchers can be interpreted as de-isolation and re-isolation moves that have contributed to predictive and explanatory improvements. In this monograph, I argue that a theoretical model, which uncovers those emotional and higher order cognitive processes causally significant for the occurrence of actual choice behavior, may bring explanatory progress not only as scope expansion but also as enhanced causal articulation and penetration.

The current chapter further explores the potential explanatory gains of building a theoretical perspective that describes those mental processes (physically executed by particular brain structures and information processing activities) that may participate in the (causal) chain of events that give rise to manifest choice behavior. To be more

specific, it tentatively provides a theoretical model of decision-making that yields a (genuinely causal) explanation of choice behavior at two levels of analysis, the mental (i.e. cognitive or algorithmic) and neural (physical brain or hardware).

The explanatory model proposed in this chapter is here interpreted as an instance of re-isolation consistent with Simon's (1992b) advice that complex cognitive phenomena like human thinking and decision-making may require a multi-level explanation. It comes to challenge the conventional economic explanation of choice behavior that is centered on a compact set of (binary) preference axioms and probability rules. The suggested two-level account is built on the working hypothesis that some neural structures and brain processing activities comprise the physical substrate of an individual's capacity to mobilize and coordinate emotional and higher-order cognitive processes with major roles in the (causal) production of important behavior patterns (some of which are commonly taken as choice anomalies).

One of the main contributions of the present chapter is to show that an account that goes one level of analysis down is worth developing since it contributes to an improved understanding of those emotional and higher order cognitive processes that participate in the causal chain of events that give rise to economically relevant patterns of choice behavior. In so doing, the current work offers an explanatory scheme that purports to fill in a gap that still remains in contemporary behavioral economic analysis – it is not clear how emotion, higher order cognition and manifest choice behavior are related in a complex economic world of causal (inter)dependencies. To organize the discussion, the remainder of the chapter is divided into 5 sections.

Section 2 analyzes features of Simon's account of bounded rationality that serve as basis for a behavioral model of choice that purports to improve the explanatory abilities of economic analysis. There is also a brief discussion of certain strengths and shortcomings of the basic theoretical model of boundedly rational choice. Section 3 poses a variant of Simon's account that incorporates ideas from neuroscience and evolutionary psychology in search of those (causally) relevant states and processes to the occurrence of choice

behavior in the real world. Based on a two-level description of the decision machinery, section 4 describes the significant ways in which emotions and feelings give rise to judgment and decision-making behavior. In Section 5, some implications of an explanation at the levels of the mind and brain are assessed. Section 6 synthesizes the overall argument and conclusions follow.

2. On Simon's model of bounded rationality

Simon's theoretical model of bounded rationality challenges the mainstream economic explanation of choice behavior in terms of a compact set of utility axioms and probability rules. Provided that his behavioral account is viewed as a reference point to further developments of models of choice that employ reformist theorizing strategies, it may be useful to discuss some of its important features and possible limitations.

Contra the neoclassical account of behavior grounded on the empirical hypothesis of expected utility maximization (e.g. Friedman and Savage 1948), Simon argued that actual decision-makers have limited computational capabilities to make optimal search of alternatives, evaluation of prospects and selection of options. In his perspective, actual choice behavior is brought about by a sequence of heuristics or rules of the thumb that boundedly rational agents employ to make judgments and decisions in the (complex) real world (Simon 1959, p.272). In the present work I interpret Simon's model of bounded rationality as a re-isolation move (i.e. it replaces the expected utility maximization principle with a series of heuristics and mental heuristics that do the explaining of manifest behavior). His major contribution is to offer an explanation of behavior that unveils mental processes (mental shortcuts, rules of thumb, routines) by which actual people can make probability judgments, estimate outcomes of salient decision options and establish a criterion to pick up a satisfactory course of action.

2.1 Building blocks of the standard behavioral model

Simon's behavioral model is premised on the idea that specialized heuristics elicit and regulate specialized cognitive processes with major roles in the production of choice behavior. He assumes that manifest behavior is shaped by a sequence of three types of

heuristics called (a) attention-focusing heuristics, (b) information search heuristics, and (c) stopping search heuristics or choice criterion heuristics (Simon 1983, 1991, 1997).

Attention-focusing heuristics mobilize specialized mental activities through which boundedly rational agents can explore information within the environment about the problem situation, make judgments about the choice task at hand and quickly evaluate its priority on the agenda for action. Search heuristics activate specialized cognitive processes (e.g. learning and memory) underlying individual's capacity to draw inferences about available choice alternatives. Finally, stopping search heuristics coordinate cognitively demanding activities just like planning and complex evaluations about the expected outcomes of salient choice options whereby the agent can select an alternative that meets his or her perceived goal priorities and concerns. 'Satisficing' is an example of stopping search heuristic that informs individuals to pick out the first alternative that meets (or surpasses) their aspiration levels. [65, 66]

One of the strengths of Simon's model of choice (relative to the expected utility approach) is that it provides a description of a sequence of cognitive processes and heuristics underlying actual behavior. The empirical literature confirms Simon's hypothesis that boundedly rational agents employ heuristics to make judgments and decisions in a wide range of phenomena that remained unexplained by standard economic analysis (e.g. Kahneman, Slovic and Tversky 1982; Gigerenzer and Selten 2001). This may partly explain why his treatment of bounded rationality became very popular among economists, psychologists and other critics of the neoclassical economic framework (Hogarth and Reder 1986, Colinsk 1996, Camerer 1998, Rabin 1998, Rubinstein 1998, Elster 1983).

2.2 Some challenges to the model of bounded rationality

Yet Simon's account of boundedly rational choice failed to deal with some difficulties stemming from economists' lack of attention to the philosophical issues of intentionality

[65] Aspiration levels can be adjusted to the time spent with information search. See Gigerenzer (2001).
[66] Gigerenzer and his collaborators from Max Planck Institute Berlin study other choice heuristics, such as 'Take the Best' (Gigerenzer, Todd and the ABC research group 1999, Gigerenzer and Selten 2001).

and causality. The model of bounded rationality assumes that individuals rely on specialized heuristics to make judgments and choices. Note that this presupposes a great deal of intentionality, i.e., the agent's desires and beliefs guide his or her usage of particular heuristics and determine a selected course of action. One possible source of objection to Simon's theoretical account is that it fails to accommodate patterns of choice behavior resulting from certain inferences and choices that are made on an automatic and non-deliberate basis. Another potential limitation of the basic model of boundedly rational choice concerns the difficulty in specifying how heuristics or mental shortcuts and the causes of behavior are related in the economic world. Putting the concern somewhat differently, Simon's theoretical account only assumes that boundedly rational agents make judgments and choices based on heuristics; thus there is no systematic description of how exactly these mental shortcuts and routines generate behavior patterns. A third, though related, source of objection to the conventional behavioral model of choice has to do with the poor conceptual links made with folk psychology. Just like the neoclassical approach to choice behavior, Simon's account makes reference to folk psychological terms (beliefs, preferences and goal aspirations) without specifying how they are related in a complex world of causal dependencies. To circumvent the above problems that seem to constrain further explanatory progress as enhanced causal articulation or penetration, I propose a variant of Simon's theoretical model of choice that draws on insights from contemporary cognitive (evolutionary) psychology and neurosciences. The idea is to identify those emotional and higher order cognitive processes with major roles in (causal) production of human decision-making behavior (Simon 1997, p.331).[67]

3. Incorporating neuroscience into behavioral analysis

This section outlines a theoretical model of boundedly rational choice that is premised on the idea that that contemporary neurosciences, behavioral decision research and cognitive psychology can inform economists about what processes may be causally significant for

[67] Perhaps Simon anticipated that behavioral decision research would converge with the study of information processing by the brain. He claims that a revolution in cognitive psychology needs to offer more details of "the connections of… information processes to the next level of theory – neural mechanisms that store information in the brain and execute information processes" (Simon 1979, p. XI).

actual patterns of choice behavior (including those regarded as anomalies).[68] The proposed variant of Simon's account is based on the hypothesis that specialized neural structures and brain processing activities physically realize some mental states and processes by which decision-making behavior comes about. It suggests that tokens of mental phenomena (e.g. states, events, processes and powers) underlying human thinking and decision-making are at bottom physical phenomena (brain processing activities), which are in turn causally sufficient to the occurrence of manifest choice behavior. Then, it implies the vision that the information-processing brain can be equated with 'what the mind does'. [69, 70]

One important feature of the suggested explanatory account is its commitment to a moderate physicalist standpoint. Physicalism is a philosophical thesis with general claims about the nature of the world (Kim 2000 and Carruthers 2004). Moderate physicalism is used here to refer to a doctrine that mental events and processes supervene on physical brain events and processes. Later in this chapter I show that endorsing this physicalist interpretation does not deny the possibility that different neural activities and processes may execute a particular mental process causally relevant to the occurrence of actual choice behavior. Furthermore, the suggested explanatory scheme does not imply a commitment to an eliminative materialist thesis (i.e. doctrine holding the view that all properties and states of the mind can be reduced to neural properties or processes).

My schematic account of decision-making behavior is based on David Marr's (1982) thesis that genuine understanding of cognitive behavior requires an explanation at different (though complementary) levels of description (pp.24-27). Marr's information-processing approach to cognitive behavior specifies three levels of analysis: hardware, algorithmic, and computational levels. In the context of an explanation of decision-

[68] For ideas on how brain studies can improve economic analysis, see Damasio (1994), Adolphs and Damasio (2001), LeDoux (1998 and 2002), Camerer, Loewenstein and Rabin (2004) and Camerer, Loewenstein and Prelec (2005).
[69] Similarly, LeDoux points out that "the things we describe in mental terms are in fact processes going on in the brain" (2002, p.195).
[70] The discussion of the implications of an explanation grounded on the physicalist thesis of mind-brain identity is left to the end of this chapter. See Carruthers (2004) for a careful discussion of why a materialist approach is consistent with beliefs about the irreducibility of certain mental events.

making behavior, the first (hardware level) specifies brain structures, states, and information-processing activities that physically realize emotional and cognitive processes underlying manifest behavior. The second (algorithmic level) describes informational inputs, mental events among heuristic processes that trigger behavioral outputs. The computational level involves a full mapping of major mental and physical brain events to domain-specific adaptive information-processing problems (that they have been designed to resolve in ancient environments). The hardware and the algorithmic levels of description refer to the proximate causes of behavior (i.e. mental and/ or physical brain events and processes that bring about manifest behavior), whereas the computational level provides an ultimate-level explanation (genetic fitness accounts governed by natural selection and transmission mechanisms via frequency dependent dynamics). [71] Even though the theoretical model that I shall shortly propose is inspired by Marr's information-processing approach to behavior, it does not offer a description of the complex system of human decision-making at the computational level. Rather, the proposed explanatory scheme focuses on the description of the decision machinery at algorithmic and hardware levels. This is because its major explanatory task is to expose the triggering forces behind economically relevant patterns of choice behavior.

In a nutshell, the explanatory account to be developed in this chapter purports to offer a description of the machinery for human decision-making behavior on two complementary levels of analysis, the mental and the neural. It goes one level of analysis down in search of those events, processes and states with major roles in manifest choice behavior in the domains of intertemporal choice, decision under risk and prosocial choice. The described decision machinery amounts to a complex system of specialized brain structures and specific information processing activities that physically execute specialized mental activities and processes that are causally relevant to the occurrence of choice behavior. A two-level account assumes that the components of the decision-making mechanism are organized in a decomposable and non-random fashion. They constitute three stages of information processing, respectively called detection stage, selective search stage and selection stage.

[71] See Mayr (1988), for a thorough analysis of proximate and ultimate causes.

The novelty of the current explanatory account is that it shows that certain emotional processes and states play important roles in the operation at all three hypothesized information-processing stages, whereas previous models fail to expose the more of less subtle ways in which emotions and feelings participate in the causal chain of events that produce actual decision-making behavior.

The first stage of processing – the detection stage - consists of specialized brain structures and information processing activities that physically execute heuristics among other mental processes that enable individuals to make quick identification and appraisal of a decision task posed by the (natural or social) environment. Note that this first stage of the decision machinery refers to the initial cognitive events and processes that Simon investigated by means of the attention-focusing heuristic. The informational outputs of the detection stage - cue-based inferences about a decision problem situation and its significance – are thought to activate the operation of the selective search stage.

The selective search stage is made up of specialized neural structures and information processing activities that serve to mobilize and to regulate mental processes as well as heuristics whereby the agent can filter information within the environment and encounter alternative choice options. This second processing stage requires a detailed analysis of those mental events and activities that Simon's classifies as search for choice alternatives. The results of the selective search stage – inferences about available courses of action – activate the working of the affect-driven selection stage.

The third processing stage from the decision machinery is constituted by particular brain structures and information-processing activities that physically realize heuristics and other mental activities, whereby a boundedly rational actor can draw inferences about probabilities and hedonic outcomes of available prospects and opt for a salient option that meets her or his perceived goal priorities, aspirations and concerns. This final stage of the machinery for decision-making behavior offers a two-level analysis of how the agent ends information search and establishes a choice criterion. Simon had something similar

in mind when his account at the algorithmic level described stopping search and agent's establishment of a simple choice criterion.

3.1 The detection stage and initialization of the decision machinery

The operation of a decision-making mechanism depends upon the agent's capacity to identify an incoming stimulus and appraise the nature and significance of the problem situation that it represents. Although these mental activities are cognitively demanding, normal individuals often make fast and effective judgments in this type of situation. This is partly because there are certain brain structures that enable individuals to exercise their ability to rely on simple heuristics (e.g. attention-focusing heuristics) and coordinate mental events and processes (e.g. sensory-processing and conceptual representation of an incoming stimulus) important to the occurrence of quick (and effective) appraisals about inferential and choice tasks posed by environments. In other words, activation of certain brain structures and areas allows the agent to seek satisfactory information about statistical regularities within the environment and to find a solution to the frame problem. The latter refers to the agent's (cognitive) difficulty with focusing attention on specific pieces of information within the environment that allow her or him to mentally represent the nature and/or importance of a choice task (Ketelaar and Todd 2000). In this work, I assume that three main brain structures play major roles in executing mental events by which an individual can draw inferences about a problem situation and the ecological significance of the task it represents: the sensory thalamus, the sensory neocortex and the amygdala.[72]

The thalamus is a subcortical portion of the brain responsible for sensory processing of information. This is largely due to its capacity to receive messages from sensory receptors (e.g. eyes, ears, skin) and mobilize processing activities that produce a quick

[72] Bear in mind that the tentatively proposed description of the decision machinery at the brain level relies on simplifications and omissions. I acknowledge that other brain structures (e.g. hypothalamus, basal forebrain) may operate in parallel to activate the detection stage. Yet I focus on the study of how the thalamus, the amygdala and the sensory neocortex may guide the first processing stage due to neuroscientific evidence showing that these three brain structures play major roles in regulating activities that enable individuals to exercise their mental ability to identify and evaluate ecologically important tasks (LeDoux 1996 and 2002, Adolphs and Damasio 2001, for instance).

(though crude) representation of an incoming stimulus. The outputs of this cue-based sensory processing by the thalamus are transmitted to the cortex.

The sensory cortex is a portion of the neocortex that is associated with the individual capability of a thorough representation of an incoming stimulus. This is partly due to its capacity to coordinate processing activities that draw upon information about sensory and conceptual inputs. The amygdala, a tiny almond-shaped subcortical structure linking the thalamus and sensory cortex, is located deep in the temporal lobe. It is responsible for processing activities that enable the individual to identify a potentially harmful or dangerous problem situation, assess its nature and ecological significance, and even mobilize his or her cognitive resources to respond adaptively (Damasio 1994 and LeDoux 1996/1998). This is partly so because the amygdala receives information about an incoming stimulus directly from the sensory thalamus and indirectly through cortical mediation. Neuroscientists seem to suggest that the amygdala's links with the thalamus and sensory cortex constitute two main sensory processing trajectories responsible for those mental activities underlying the detection stage (i.e. initialization of the decision machinery).[73] Neuroscientists call them the thalamo-amygdala and thalamo-cortex-amydgala routes (ibid).

According to Joseph LeDoux, the thalamo-amygdala trajectory is shorter and entails faster transmission of sensory messages (perceptual inputs) than the thalamo-cortex-amygdala pathway. The thalamus receives sensory information from external receptors (eyes, ears, etc) and relays outputs directly to the amygdala. These subcortical activities provide the neural basis of the agent's ability to identify an incoming stimulus and quickly evaluate the decision problem involved. The cortical road (i.e. thalamo-cortex trajectory), in its turn, offers the neural underpinnings of cognitive activities involved in a rigorous sensory and conceptual representation of a problem situation. As he put it,

> Although the thalamic system cannot make fine distinctions, it has an important advantage over the cortical input pathway to the amygdala. That advantage is time… The thalamic pathway is thus faster. It cannot tell the amygdala exactly what is there, but can provide a fast signal that warns

[73] They offer evidence that individuals with impaired amygdala have more difficulty with identifying an aversive stimulus and to draw quick inferences about a choice task than normal patients, even when their higher order cognitive capabilities are intact. For details, see Damasio (1994) and Adolphs and Damasio (2001).

> that something dangerous may be there. It is a quick and dirty processing system (LeDoux 1996/1998, p.163)
>
> […] The information received from the thalamus is unfiltered and biased toward evoking responses. The cortex's job is to prevent the inappropriate response rather than to produce the appropriate one… From the point of view of survival, it is better to respond to potentially dangerous events as they were in fact the real thing than to fail to respond (p.165).

The passage emphasizes that the functioning of both brain routes are important for the individual's capacity to make cue-based judgments about important choice tasks. In LeDoux's example, processing information about a dangerous stimulus may activate fear-driven heuristics by which agents mobilize and coordinate perceptual, atttentional, learning and memory processes in order to identify and appraise immediately an important choice task.

If this is so, the thalamus, sensory cortex and amygdala may be three major structures operating jointly to process information (often in parallel) that helps agents focus their attention on details of recurrently important challenges posed by the environment and compile estimates of possible ways of responding adaptively.

Diagram 1 provides a schematic description of major components of the detection stage that coordinate processing activities underlying the agent's mental capacity to identify a problem situation and assess the nature and/or significance of the choice task that it represents.

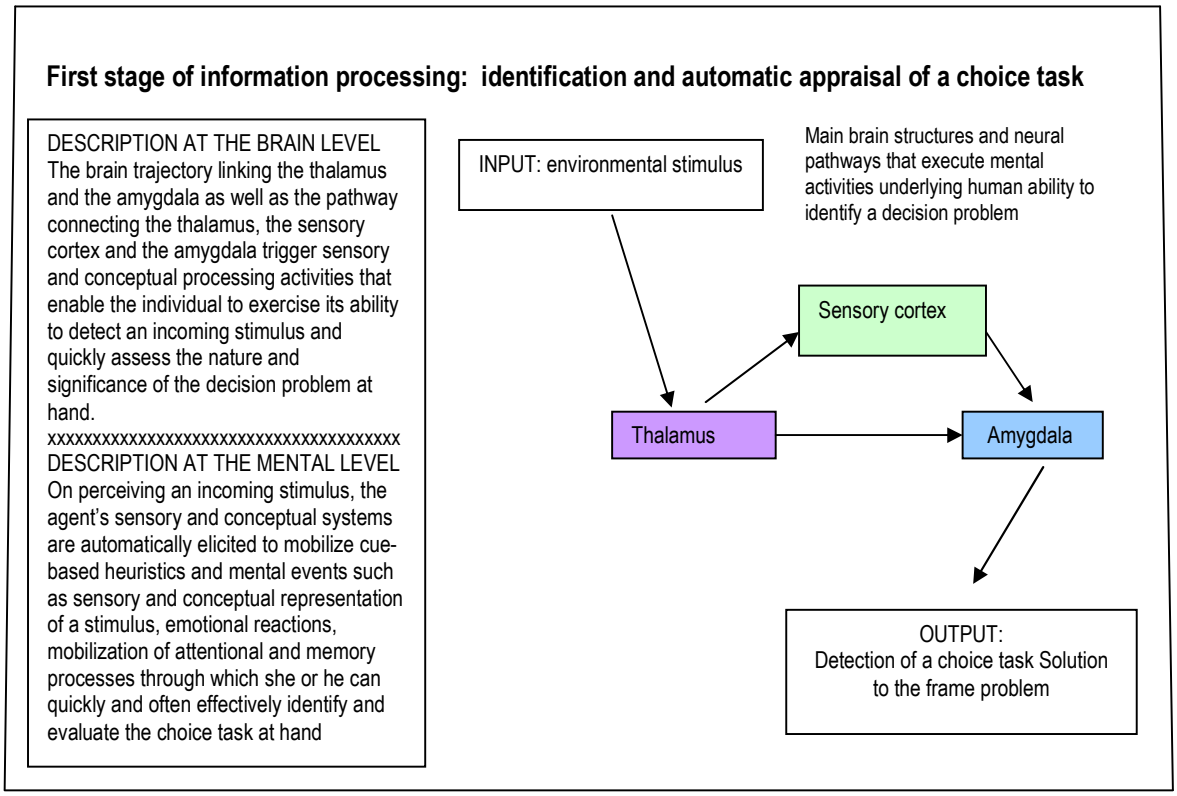

DIAGRAM 4.1

Detection stage: description of how specialized brain events mobilize mental activities underlying individual's ability to detect an incoming stimulus and evaluate the nature (and significance) of the choice task it represents. The colored boxes and arrows refer to specialized brain structures and neural pathways that physically execute mental activities (cognitive and affective processes) underlying the individual's ability to identify and assess a choice task.

Thus specialized brain structures (sensory thalamus, amygdala and sensory cortex) may play major roles in the mobilization of processing activities significant for the detection of an important choice task. They realize physically mental events and processes whereby individuals can make perceptual and conceptual representations of incoming stimuli and identify the choice tasks they yield. The outputs from the detection stage serve as input data for the second stage of processing within the decision- making system.

3.2 The selective search stage and the task of seeking choice alternatives

A great deal of decision-making probably consists of seeking possible ways of resolving a problem situation. Simon (1997) posed the issue judiciously,

> If we look at the time allocations of economic actors… we find that the largest fraction of the decision-making time is spent in searching for possible courses of action and evaluating them (i.e. estimating

> their consequences). Much less time and effort is spent in making final choices, once the alternatives have been generated and their consequences examined (pp. 291-292)

Given that actual decision-makers do not have unlimited time, energy and even cognitive capabilities to fully explore information within the environment about the choice task at hand, their search for choice options happens on a selective basis. For this reason, I call the second processing stage within the decision machinery 'selective search'.

Supplementing the basic model of boundedly rational choice that describes specialized heuristics that enable agents to seek alternative choice options, the theoretical account proposed here includes a description of those brain events and processes by which the agent can exercise his or her mental capacity to draw quick and effective inferences about salient choice alternatives. Neuroscientific studies indicate that four specialized brain structures - the prefrontal cortex, hippocampus, amygdala and basal ganglia – serve in central roles in instantiating mental activities (heuristics, memory and learning processes) that in turn enable the agent to encounter potentially satisfactory ways of coping with the choice task at hand (LeDoux 1996/1998 and 2002).

The prefrontal cortex is a portion of the neocortex responsible for activating and coordinating those processing activities that underlie higher order cognition and executive functions of the mind, such as planning and decision-making.[74] According to neuroscientists, the prefrontal cortex receives information from many brain structures and bodily systems. Through these connections, the prefrontal cortex can mobilize and coordinate brain events and informational processes by which the individual can exercise his or her ability to draw inferences about alternative ways of dealing with a choice task. The lateral region of the prefrontal cortex is responsible for realization of mental states and activities known as working memory (LeDoux 2002). The latter refers to cognitive and emotional processes and cue-based heuristics through which the agent can integrate information about a current choice task with her or his long-term knowledge of important

[74] For details, see Damasio (1994, 1999) and LeDoux (1996/1998, 2002).

decision problems (and how they were resolved).[75] Neuroscientists offer evidence that the lateral prefrontal cortex offers the physical substrate for the agent's capacity to learn from its own and others' previous experiences (Adolphs and Damasio 2001).[76]

The hippocampus is a subcortical structure that undertakes processing activities responsible for the agent's mental capacities such as episodic, declarative, contextual, spatial memory processes and learning (Squire and Fanselow 2000). The hippocampus seems to physically execute specialized mental phenomena through which individuals can store (or retrieve) information of earlier situations or events, the circumstances in which they happened, and how they were resolved. In this case, the functioning of the hippocampus plays an important role in guiding the quality of agent's estimations of alternative choice options.[77]

The amygdala also plays an important role in the stage of selective search. Through its connections with the hippocampus and prefrontal cortex, the amygdala mobilizes and regulates processing activities responsible for encoding and retrieving information about harmful situations (and how to resolve them). If this is so, the amygdala can be viewed as one of the physical foundations of mental processes (e.g. attention, learning and memory) that guide quick and effective judgments about possible ways of dealing with a harmful situation.[78] The basal ganglia are the "amygdala counterpart" for encoding and retrieving information about a rewarding stimulus or beneficial problem situation. They activate and

[75] The concept of 'knowledge' used here is broad, since it refers to propositional knowledge (justified true beliefs), procedural or skill knowledge (knowing how and what) and even acquaintance knowledge (knowing who).

[76] Based on experimental studies, Adolphs and Damasio (2001) claim that impaired prefrontal cortex patients have more difficulty than normal subjects in drawing effective inferences about available options (and hedonic consequences associated with them). They suggest that impaired prefrontal cortex agents have limited ability to use earlier experiences vis-à-vis current perceived states as sources of inferences about choice alternatives.

[77] Individuals with an impaired hippocampus have difficulty in estimating future outcomes on the basis of previous experiences (ibid).

[78] There is evidence that individuals with amygdala lesions have deficits in encoding, retrieving and learning information about potentially harmful or dangerous stimuli (LaBar and LeDoux 2003, p.59). Based on experimental studies, they find results that confirm the hypothesis that amygdala-impaired patients often make poorer judgments and decisions than normal individuals (ibid).

regulate processing activities underlying the mental ability to rely on cue-based learning and memory processes to estimate salient choice options (Lane and Nadel 2002; Davidson, Scherer and Goldsmith 2003). Diagram 2 sketches the selective search stage, and suggests what specialized brain structures may do to physically execute mental activities that enable individuals to seek and find alternative courses of action.

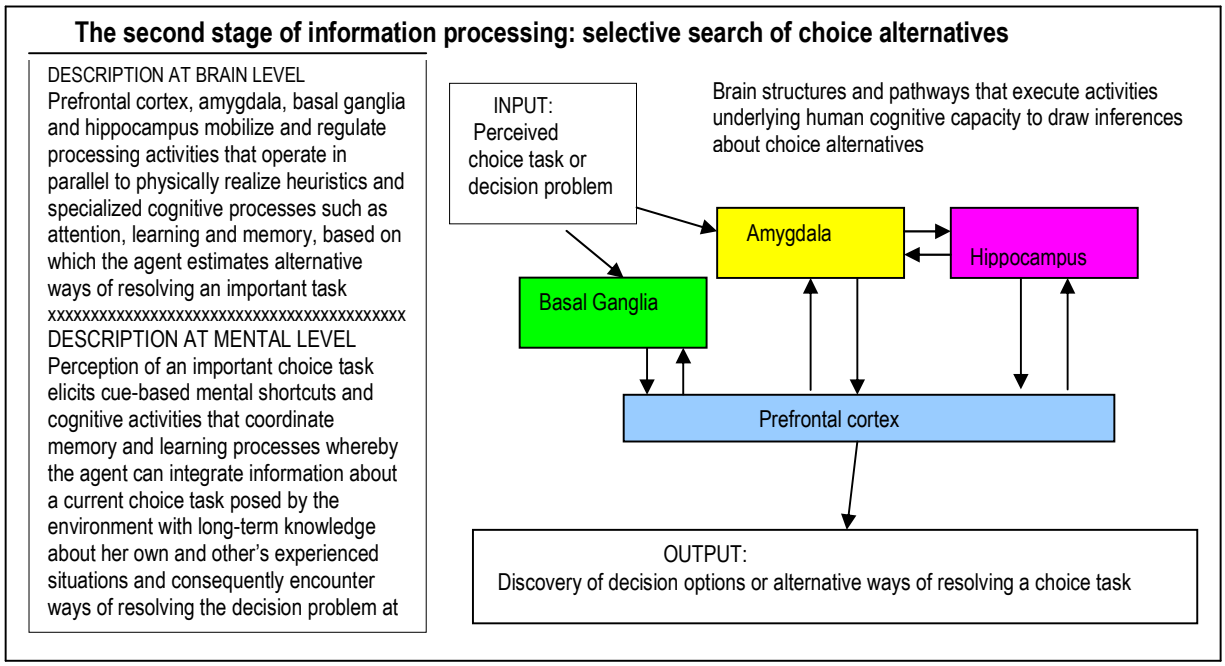

DIAGRAM 4.2

Description of the search stage in which specialized brain events regulate and coordinate processing activities that physically execute specialized cognitive processes by which the agent can exercise her or his mental capacity to draw inferences about satisfactory choice options.

In short, specific brain structures like the prefrontal cortex, amygdala, basal ganglia and hippocampus may work in parallel to regulate processing activities that physically execute mental processes (like memory and learning processes) through which the individual can allocate her or his limited computational resources to make judgments about available courses of action. The output of this second stage of processing – discovery of choice alternatives – serves as input for the activation of the next stage of the decision system.

3.3 The affect-driven selection stage and determination of choice (behavior)

In this last processing stage within the decision machinery, specialized neural structures and activities work together to instantiate physically those cognitive and affective processes underlying the agent's capacity to assess the outcomes of salient choice options and to pick out one that meets her or his goal priorities and aspirations. [79]

The description of the machinery for human decisions proposed here ascribes a central role to feeling states at the selection stage. By virtue of their capacity to alter the agent's perception of her or his goal priorities, aspirations and concerns, feelings guide individual selection of a course of action (to the exclusion of others). Based on insights from neurosciences, feelings are thought to play major roles in the operation of the decision machinery to the extent that they serve the function of body-related signals (i.e. somatic markers) about rewarding and punishing consequences of alternatives (Damasio 1994). Provided that feelings involve positive or negative hedonic values they are here interpreted as mental states that serve as reliable cues about the courses of action that one ought to pursue and the ones that need to be avoided, respectively (ibid).[80] Three brain structures seem to play major roles in processing activities by which individuals can exercise the mental capacity to opt for a particular course of action/ behavior pattern. They are the prefrontal cortex, the amygdala, and the nucleus accumbens.

At this stage of the decision machinery, the prefrontal cortex serves to regulate information processing activities that instantiate higher order cognition, such as planning, means-end reasoning, cost-benefit weighing and choice-making. This is partly so because the prefrontal cortex can coordinate pieces of information coming from different brain areas and overall body (LeDoux 2002, p.180). According to neuroscientists, two portions of the prefrontal cortex - the orbital cortex and the cingulate cortex – may play central roles in executing the agent's capability for quick and effective choices. These brain

[79] Economists do not deny that affect has a role in choice behavior to the extent that it shapes human preferences. It seems to me that many (affect-based) preferences are left inscrutable in the black box of consumer theory due to the strong influence of revealed preference theory and of its extensions (discounted utility theory, expected utility theory and subjective expected utility theory) on their work. Only recently economists have pursued a process-description strategy to investigate the foundations and structure of agents' preferences. See Camerer, Loewenstein and Prelec (2005) for a discussion of this issue.

[80] See Cosmides and Tooby (2000) for a discussion about the evolutionary functions of emotions.

structures seem to be responsible for encoding, decoding and updating information about important situations as well as about reward or punishment associated with alternative courses of action.[81] More specifically, the orbital cortex coordinates activities that yield individuals the capacity to process information about rewarding and punishing prospects.[82] The anterior cingulate cortex mobilizes processing activities that inform the agent how certain emotional states (feelings) approximate her or his perceived aspirations and goal priorities (Ochsner and Feldman Barret 2001).[83]

The amygdala is an important structure for processing information about aversive stimuli and emotionally arousing situations recurrently posed by the environment. This is partly due to its connections with the prefrontal cortex and the hippocampus. The amygdala is responsible for the activation of processing activities that physically realize the individual's capacity to estimate the hedonic outcomes of available courses of action and to select an alternative that meets her or his perceived preferences, aspirations and so forth.[84]

The nucleus accumbens is a forebrain structure that guides processing activities with major roles in human motivation and reward ascription. This is due to its biochemical constitution. The nucleus accumbens is endowed with a stock of dopamine inputs received from a forebrain region called ventral tegmental area that constitutes the motive circuit. By virtue of its connections with brain structures responsible for sensory processing (e.g. the amygdala and the cortex) and motor control (e.g. the pallidum), the

[81] Based on experiments with brain-damaged patients and normal individuals, neuroscientists confirm the hypothesis that the quality of decision-making depends on activities performed by the prefrontal cortex (Bechara, Damasio, Damasio and Anderson 1994, Bechara, Tranel, Damasio and Damasio 1996).
[82] According to Damasio and his associates, patients with impaired orbital cortex have lowered capacity to make quick and clean judgments and decisions under conditions of uncertainty, conflict and limited time for deliberation (Bechara et al 1995, 1997).
[83] Individuals with damaged cingulate cortex have difficulty in estimating probabilities and hedonic values associated with available prospects. They also have lowered capacity to respond adaptively to an aversive stimulus and to discriminate good choice options from bad ones (Damasio 1994; Adolphs and Damasio 2001).
[84] Bechara et al (1999) find evidence that amygdala-impaired individuals make poorer decisions than normal individuals. They show that damage to the amygdala constrains individual's capacity to rely upon earlier experiences and currently experienced affective states (feelings) as reliable cues about rewarding and punishing options (to be pursued and avoided).

nucleus accumbens mobilizes processing activities that inform cortical regions about the rewarding or punishing consequences of salient prospects to be pursued or avoided. [85, 86]

To sum up, the constituents of the selection stage operate as follows. At the very identification of alternative courses of action, specialized structures like the prefrontal cortex, the amygdala and the nucleus accumbens may be automatically elicited to mobilize and regulate processing activities that in turn enable an individual to exercise her or his capacity to estimate the probabilities and hedonic outcomes of salient choice alternatives, to establish a criterion to select an option that is felt to satisfy or exceed perceived goal priorities and aspirations and, therefore, to exhibit a satisfactory choice behavior.[87] Given the cognitive burden of such inferential activities, actual individuals take their feeling states as cues to select a potentially rewarding prospect to the exclusion of less rewarding or punishing ones. If this is so, it makes sense to suggest that the agent's experience of emotions activate a cue-based selection heuristic (tentatively) called 'hedonic satisficing': the agent picks out the first decision option that is felt to meet or exceed her or his perceived goal priorities, preferences, concerns and so forth.[88] Diagram 4.3 summarizes the inner operations of the selection stage and how its components work to give rise to actual choice.

[85] The pallidum regulates processing activities that are productive of affective states, such as pleasure associated with the sight and taste of food.
[86] According to Mogenson, Jones and Yim (1980), the nucleus accumbens sits at the crossroads of emotion processing and behavioral movement.
[87] Notice that this statement sheds light on how the brain level is related to the mental level. Recall that specialized neural structures and processing activities physically execute certain mental activities (cognitive and affective processes) playing a significant role in generating manifest choice behavior. Indeed this idea is laden with a rather complex ontology of mental causation. The two-level account presupposes that some brain processes (denoted as information-processing activities) have to occur for the agent to be going through mental events and activities (emotional and higher order cognitive processes) playing major (causal) roles in production of the decision-making phenomenon. This implies that the vision that domain of mental phenomena supervene on the neural domain.
[88] For a similar line of argument, see Ketelaar and Clore (1997), Ketelaar and Goodie (1998) and Ketelaar and Todd (2000).

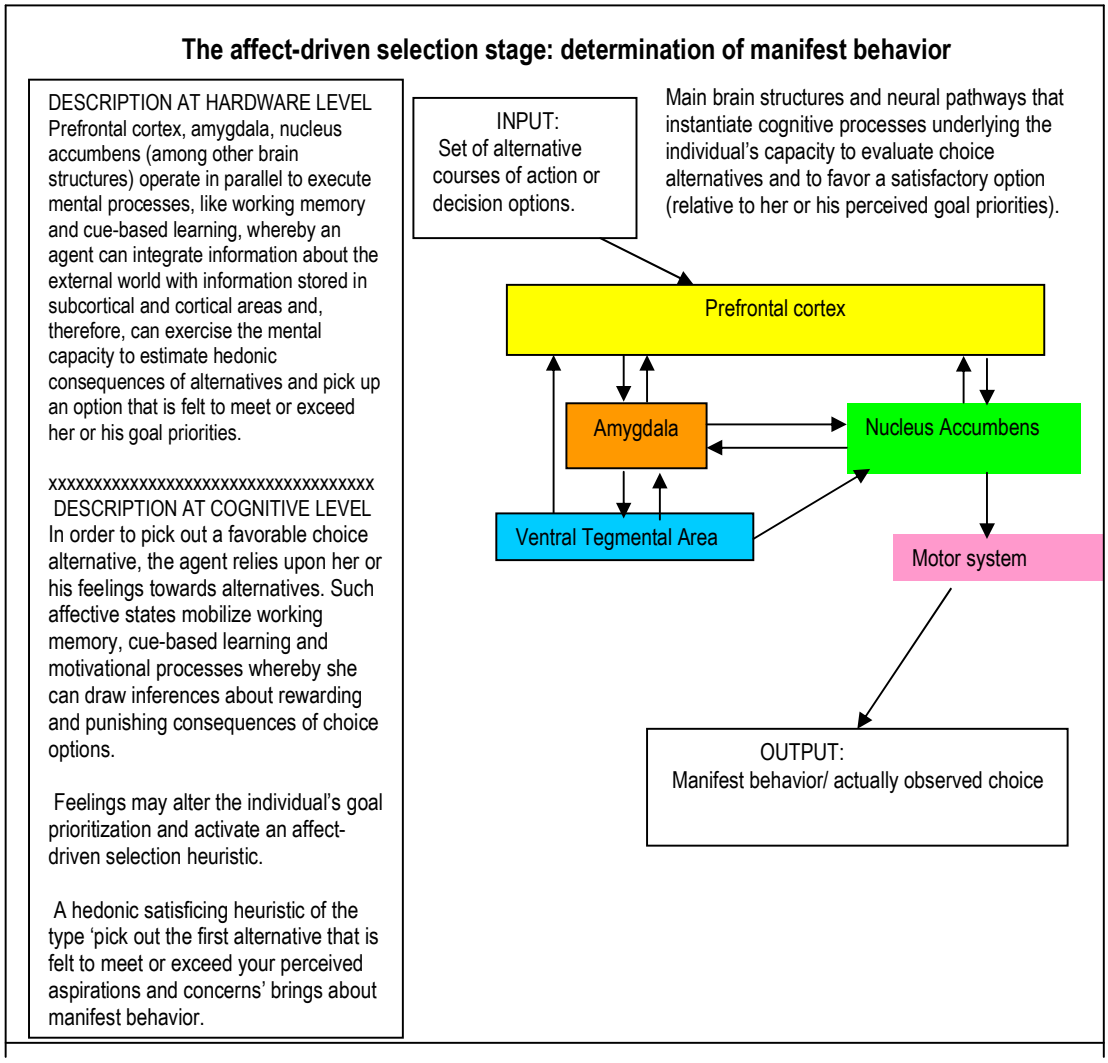

DIAGRAM 4.3

Description of the selection stage in which specialized brain structures (represented in the colored boxes) may activate and coordinate processing activities that enable individuals to assess salient options, draw inferences about their consequences and to select a course of action that meets the agent's perceived aspirations and goal priorities.

Figure 4.4 shows a picture of the brain (and its neural structures) that may interest economists attempting to open the black box of processes and mechanisms through which actual decision-making behavior comes about (see LaBar and LeDoux 2003, p.59).

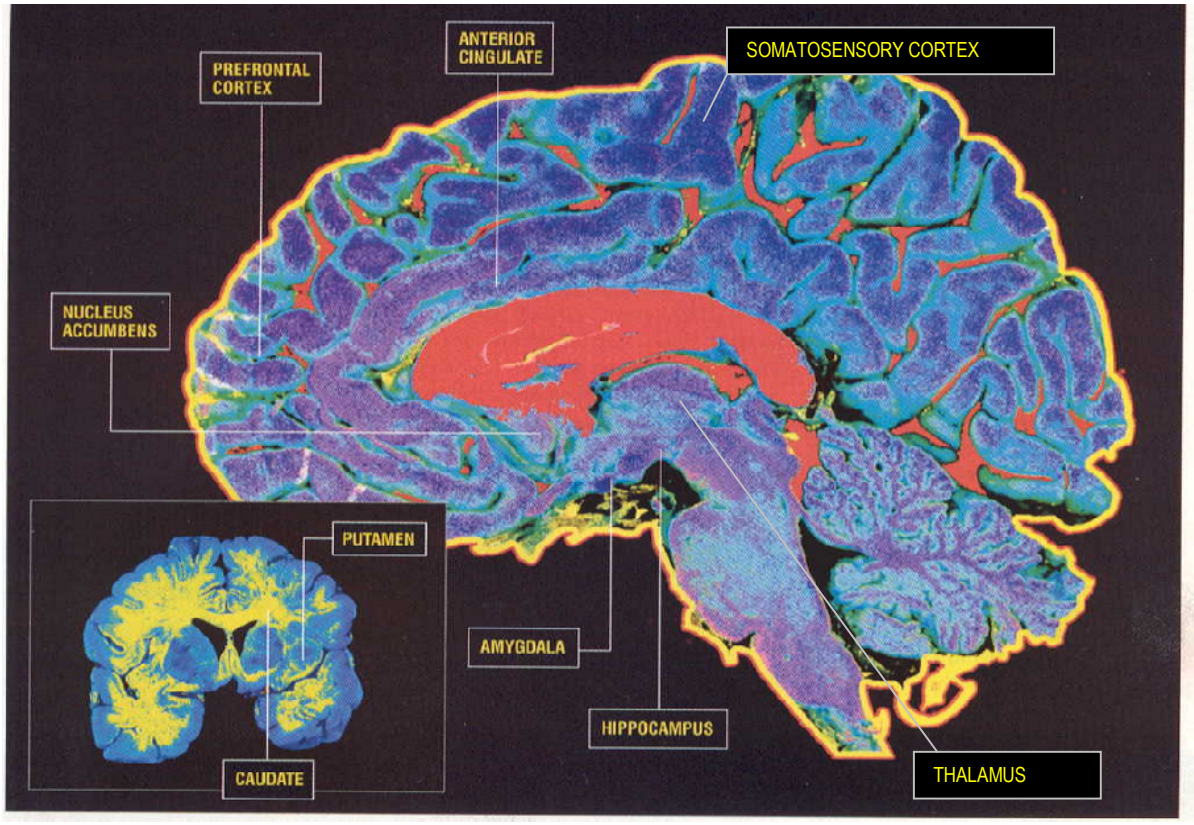

FIGURE 4.4

This slice of the brain shows neural structures that execute mental states and processes playing major roles in decision-making behavior. The human brain contains 4 lobes. In a clockwise direction (from left to right), we can identify the frontal, parietal, occipital and temporal lobes. The first is assumed to be a brain region responsible for executive functions of the mind such as planning, cognitively controlled activities, and integrating cross-brain informational inputs. The second lobe involves motor control. The occipital lobe is thought to be responsible for visual processing. The temporal lobe is involved with the mental capabilities for memory, learning, recognition and emotion.

Source: Camerer, Loewenstein and Prelec (2005, p.17)

Note that the proposed theoretical model of choice offered a description of how the decision machinery operates to produce behavior. Its distinctive trait is to provide a two-level explanation of decision-making that describes the significant role of emotions in the chain of mental events and processes that produce economically relevant patterns of choice behavior. In the next section, I go on to detail the more or less important ways in which emotions and feelings may influence the complex machinery for human decisions.

4. How emotion and feeling shape the decision machinery

Challenging the vision that emotions enter into the explanantia set of economic analysis of choice as mere arguments of the individual utility function or constraint parameters (Becker 1996, Elster 1999), the above-mentioned two-level description of the decision machinery identifies the (causally) significant roles that emotions among other affective processes play in (a) automatic detection of a problem situation posed by the natural or social environment; (b) effective search for choice alternatives that resolve the perceived choice task, and (c) quick and satisfactory selection of a course of action that corresponds to the agent's felt aspirations and concerns. Provided that the suggested theoretical model purports to yield a mechanistic explanation of decision-making behavior that offers scientific understanding of how things happen in the complex economic world, it may be useful to discuss how emotions among other affective states (e.g. feelings, moods and visceral states like hunger, thirst, sexual desire or craving for addictive goods) shape the machinery for human decisions.

4.1 Emotion and the detection stage

Evolutionary psychologists tell us that basic emotions such as fear play the role of (hard-wired and learned) automatic detectors of threats and opportunities recurrently posed by the (social and/or natural) environment.[89] Just like a set of bodily reactions and of changes in cognitive processing, emotional processes activate cue-based inferences about the significance of an incoming stimulus and the adaptive problem situation that it represents (Cosmides and Tooby 2000, Ketelaar and Todd 2001). Empirical findings by psychologists and neuroscientists portray an emotion as a specialized computational device or information-processing system that facilitates quick and effective identification and appraisal of an adaptive choice task, and thus may exert significant influence on the initial processing stage within the decision machinery (see also Tooby and Cosmides 1990).

[89] Other examples of basic emotions are anger, disgust, sadness and surprise. For details on the discussion about what it is that is basic about these emotions, see Plutchik (1980) and Ortony and Turner (1990).

Based on my description of the detection stage, it could be suggested that specialized emotions activate simple (i.e. cue-based) heuristics that reallocate the agent's attention and perception (and other limited cognitive resources) in order to yield quick and effective assessments of an important problem situation. In developing a similar line of argument, Simon (1983, p.27) claims that one of the crucial roles emotions play in decision-making is to enable the agent to focus her or his attention on an important choice task and find a quick and effective solution to the frame problem. For instance, the emotion of fear influences the individual's ability to identify a recurrent danger (e.g. the risk of being stalked and killed by a hungry predator) and to prepare her or his mental capabilities for responding adaptively (e.g. fight, freeze or flight).[90] More recently, Ketelaar and Todd (2000) suggest:

> Emotions… can perform the role of internal alarm systems, alerting us to adaptive dangers and benefits by inferring the distal future from one's current [and previously experienced] circumstances… In this way, emotions can help the computationally limited human mind to circumvent the pitfalls of the frame problem by determining which information to attend in the first place (p.204)

Hence, emotions offer reliable environmental cues that guide quick (automatic) appraisals of recurrent challenges posed to an individual and contribute to the discovery of a quick (and often effective) solution to the frame problem. Neuroscientific research seems to confirm this line of argument (LeDoux 1998 and 2002, Adolphs and Damasio 2001). For instance, there is evidence that some brain structures responsible for processing activities playing major roles in the detection stage (e.g. amygdala, thalamus and sensory cortex) also refer to key structures in emotion processing (LeDoux 1996/1998). Damasio and colleagues rely on experimental results to suggest that individuals with impaired amygdala become emotionally flat and therefore tend to decide less advantageously than normal agents. Recall that the amygdala is thought to mobilize processing activities that offer the physical underpinnings of the agent's capacity to detect automatically an aversive stimulus and to react adaptively to it (Bechara et al 1997 and Adolphs and Damasio 2001).

[90] More details on this example are found in Cosmides and Tooby (2000).

4.2 Emotion and selective search

Many emotions provide a frame for the agent's mind in shaping her or his cognitive facilities (e.g. attention, memory and learning) in identifying alternative ways of resolving a perceived choice task. If this is so, it makes sense to assume that they influence the selective search stage. Cognitive psychologists suggest that particular emotions activate cue-based heuristics and inferential processes by which an individual can draw quick inferences about available courses of action (vis-à-vis her own or others' previous experiences). In addition, emotions are accompanied by certain conscious affective states (e.g. feelings) that serve to narrow the individual's salient choice alternatives (Damasio 1994). Therefore emotions may resemble evolved computational devices (i.e. domain-specific information-processing systems) designed to inform the individual about available courses of actions within her or his choice set (Todd and Ketelaar 2000, p.194).

4.3 Feeling and selecting of a satisfactory course of action

Determination of choice behavior depends on various cognitively demanding processes through which the agent estimates the expected outcomes (payoffs) of salient choice alternatives and evaluates what course of action will satisfy her or his perceived goal priorities and concerns. In the above described theoretical perspective, feeling states play a crucial role in mobilizing processing activities whereby decision makers can exercise their capacities to select a favorable choice option.

Feeling states (physically instantiated by particular brain structures, such as the amygdala and prefrontal cortex) offer cues about the hedonic consequences of salient choice alternatives. In addition, feelings shape agent's perception of available prospects and their consequences and motivate the use of a hedonic satisficing procedure that picks out the first choice alternative that is felt to meet the agent's perceived aspirations and concerns. This is largely so because feeling states are accompanied by positive or negative hedonic values that serve as cue-based sources of inference about choice alternatives and their beneficial or harmful consequences. This presumption is in tune with Damasio's hypothesis that feelings are somatic markers. In his interpretation,

> The emotional signal marks options and outcomes with a positive or negative signal that narrows the decision-making space and increases the possibility that the action will conform to past experience (2003, p.148)

The underlying idea is that some positive feeling states attached to choice alternatives serve as body-related signals that inform the agent about options that have been associated with rewarding consequences. Negative feeling states offer cue-based sources of inferences about punishing outcomes. As a result, feeling states narrow the agent's perceived opportunity set and motivates her or him to select a favorable course of action.[91]

Since feeling states mobilize cognitive processes whereby agents can draw inferences about expected payoffs of choice options and also alter the individual's perceived goal priorities, it makes sense to suggest that they might elicit an affect-driven selection heuristic of the type "pick out the first salient alternative that is felt to satisfy (or exceed) your perceived aspirations, preferences and concerns". For that reason, it is dubbed 'hedonic satisficing heuristic'. Diagram 4 summarizes the main components of a two-level explanation of choice behavior.

[91] Quite similarly, Ketelaar and Clore (1997) argue:

"Feelings provide highly credible information about the appeal of choice alternatives, but in addition to this high road to decision-making in which affective information is included in judgments and deliberations, affect also takes the low road, serving directly as a pleasant incentive or an aversive punishment making one approach or avoid certain choices" (1997, p.379).

Constituents of a two-level account of decision-making behavior

Stages	Brain Structures	Relevant information processing	Significant cognitive processes	How emotion and feeling may influence the decision machinery
Detection and early appraisal of a choice task	Thalamus, Amygdala, Sensory Cortex	Activation and mobilization of sensory and conceptual processing activities that underlie mental events and processes by which individuals can draw inferences about an incoming stimulus and mentally represent the nature and significance of the choice task that it poses.	Certain perceptual and conceptual processes interact to coordinate informational outputs that serve as sources of cue-based inference about an incoming stimulus. They elicit specialized heuristics and mental processes like attention and memory through which a quick and often effective mental representation of the nature and significance of a choice task may be made.	Emotions serve as automatic detectors of recurrently dangerous and beneficial stimuli posed by social/ natural environments Emotions (bodily reactions) offer cues that facilitate inferences about an urgent choice task with ecological significance.
Selective search for alternatives	Prefrontal Cortex, Amygdala, Basal Ganglia Hippocampus	Elicitation of emotion and higher-order cognitive types of processing that serve to integrate information about the external world and long-term knowledge about one's own and/or other's previous situations and experiences. They constitute the neural underpinnings of agent's mental capacity to seek and find alternative courses of action.	Cue-based mental shortcuts and mental activities activate and regulate attention, memory and learning processes underlying quick and often effective inferences about alternative courses of action (or possible ways of resolving a particular choice task)	Whenever a choice task is perceived as being urgent, a specialized emotional reaction is elicited to prepare the body to concentrate cognitive resources (attention, memory, cue-based learning) on seeking alternatives that resolve an adaptive choice problem.
Affect-driven selection of a course of action	Prefrontal cortex, Amygdala, Nucleus accumbens	Higher order cognitive processing and emotion processing integrate information about the external world and the agent's internal objective and subjective world. They provide the physical substrate for an individual's ability to make plans, to estimate outcomes and probabilities of choice and to select a choice alternative that approximates her or his perceived aspirations.	Affect-driven mental heuristics and mental activities like working memory, cue-based learning and motivational processes coordinate information from the environment and long-term knowledge about previous experiences so as to enable the agent to draw quick and effective inferences about hedonic outcomes of choice prospects and to select a rewarding alternative that first satisfies or exceeds her or his perceived goal prioritization or aspiration.	Feeling states ascribed to choice alternatives are reliable cues about rewarding or punishing outcomes to be pursued or avoided. They motivate the selection of a particular decision option by altering one's perceived goal prioritization or aspiration level and they eventually activate an affect-driven selection heuristic (hedonic satisficing procedure)

DIAGRAM 4.6
Ingredients of a two-level behavioral explanation of choice behavior
The blue boxes specify the sequence of processing stages organized in a non-random decomposable fashion to explain how choice behavior comes about. The operation of the decision

machinery starts with the agent's identification of an external stimulus and conceptual representation of a choice task and ends with her or his actual selection of an affectively salient choice alternative that is felt to satisfy her or his perceived aspirations. The yellow boxes describe brain structures that execute physically cognitive processes underlying choice behavior takes place in the world. The pink boxes describe the main information processing activities involved, whereas the orange boxes suggest the ways in which emotions guide choice behavior

Thus, the above-mentioned theoretical model of choice yields an explanatory account of choice behavior in that it describes mental processes and states playing major roles in (causal) production of decision-making behavior. By describing how physical brain components of the decision machinery operate, the model improves our grasp of how choice behavior comes about in the world's structure of (causal) dependency relations. It offers a first step towards uncovering the chain of events that link the complex explanandum phenomenon of decision-making behavior to its causally significant explanantia items.

5. Implications of a two-level behavioral explanation

As an attempt to open up the black box of choice theory, the proposed variant of Simon's theoretical model offers three important advantages. First, it may offer an additional check on whether a particular theoretical account correctly isolated the explanatory factors that play major causal roles in choice behavior.[92] Second, it yields understanding of why emotions are no minor factors in the machinery for human thinking and decision-making but also indispensable ones.[93] Third, it offers a detailed description of a mechanism for behavior that gives extra intelligibility to the complex phenomenon of decision-making.

5.1 An extra check on the significance and robustness of described causal processes

One of the merits of the proposed explanatory approach is that it incorporates a physical brain (hardware) level of analysis that supplements an account of choice behavior in terms of its underlying mental processes. More precisely, the inclusion of a brain level of

[92] For an explanation at the physical level to genuinely act as an extra check of an explanation of the mental level, we need to learn more from neuroscientists how specific brain areas/states and particular mental states are related.
[93] See Damasio (2003), p. 145

description offers an additional way of investigating the (causal) significance and robustness of the isolated emotional and cognitive processes for the occurrence of actual choice behavior. In this context, studying robustness as changes in implied causal processes or mechanisms is worth studying because it might inform behavioral scientists about whether our theoretical model can tell a correct (or approximately true) causal story of the phenomenon (e.g. decision-making behavior) under study.[94]

The suggested account implies the idea that the neural level of analysis of choice behavior contributes to the development of an explanation that successfully picks out the major properties and events in our explanandum's causal history. This is at least partly so because the proposed theoretical model is committed to a particular interpretation of mind-brain (causal) dependency relations holding that mental properties and events (e.g. emotional processes, human thinking and decision-making) are supervenient on brain properties and events (e.g. amygdala, prefrontal cortex, nucleus accumbens) that are in turn causally sufficient to manifest choice behavior.[95] In this perspective, no mental properties can only have non-physical realizations (Kim 2000). Note that the above ideas square well with the behavioral economist's view that neuroscientific studies may inform economists about those mental processes relevant to the occurrence of behavior.

5.2 Further intelligibility to complex behavioral phenomena

The two-level explanatory account of decision-making behavior has the merit of offering a mechanism description that gives extra intelligibility to the complex phenomenon of decision-making. By uncovering a complex system of information-processing activities performed by specialized brain structures that execute physically mental processes underlying choice behavior, the suggested explanatory account allows for an improved

[94] For an interesting and detailed study of robustness in economic models, see Guala and Salanti's papers. They suggest that robustness is often regarded as a desirable feature and even related to the prospect of scientific progress. This is because progress involves the development of models that describe increasingly robust (causal) relations, i.e., processes or mechanisms by which systematic and non-random patterns of behavior are produced. In this case, the study of robustness (defined as changes of described causal processes) is important since it offers evidence whether our models of choice failed to provide a unique account of the phenomenon under study or to detect the correct mechanism at work (p.6).

[95] For details about the physicalist thesis and its variants, see Kim (2000) and Carruthers (2004). The physical realization perspective hypothesizes that no mental properties can only have non-physical realizations.

understanding of the places of emotion, cognition and decision-making in the causal chain of events that give rise to manifest behavior. As I tried to point out, such advantages seem to outweigh the loss of simplicity and parsimony that a two-level explanatory account of decision-making behavior involves. Of course assessing the validity of my bold claims is partly an empirical matter. For that reason, chapters 5 to 7 examine the consequences of the proposed explanatory scheme for scientific understanding of important domains of human affairs, such as intertemporal choice, decision under risk/uncertainty and prosocial choice.

5.3 Understanding of why emotions are non-negligible components of the decision machinery

The two-level theoretical model offers a description of a decision-making mechanism that specifies the more or less explicit ways in which emotions and feeling states guide the three information processing stages of the machinery that brings about actual choice behavior. It suggests that emotions serve as reliable cue-based sources of inference that enable agents to identify important problem situations and to mobilize their limited computational resources to respond adaptively to them. In this case, emotions may play non-negligible roles in the detection of important choice tasks. Additionally, the model shows that emotions guide some individual's attention, memory and learning processes by which the agent can undertake a selective search for choice alternatives. This is partly so because emotions serve the evolutionary functions of cognitive guidance and coordination (Cosmides and Tooby 2000). Lastly the two-level model implies the vision that emotions play a decisive role in the production of choice behavior because they are followed by conscious affective experiences viz. feelings that alter the agent's perception of goal prioritization and motivate her/his to opt for a particular course of action to the exclusion of others. . In this case, it challenges the standard view of emotions as mere sources of individual preferences (i.e. arguments of the utility function) or constraint parameters (Becker 1996). A two-level explanation of behavior is worth developing since it yields understanding of why emotions and feelings are central (rather than negligible) forces behind and often effective thinking and decision-making in the real world of boundedly rational agents (Damasio 2003, p.145).

5.4 Some metaphysical consequences

The two-level account of decision-making behavior implies a commitment to a physicalist position in the philosophy of mind. It presupposes that mental phenomena like choice behavior can occur because individuals are equipped with certain physical brain structures with innate properties and acquired abilities enabling them to exercise particular mental capacities like thinking and decision-making. If this is so, the physicalist perspective adopted in this chapter also endorses the notion of supervenience. The latter suggests the existence of some covariation as well as some causal dependency relations between properties, states and processes of the mental realm (e.g. emotion, higher order cognition, thinking and decision-making) and those of the physical brain domain (e.g. information processing activities mobilized by amygdala and cortex trajectories).[96]

Furthermore, a two-level account of decision-making behavior subscribes to the physicalist idea that mental phenomena are at bottom physical phenomenon. It hypothesizes that the thinking mind can be equated with information-processing by the brain. This implies a commitment to a doctrine called mind-brain identity thesis (Kim 2000 and Carruthers 2004).[97] There are two versions of the mind-brain identity thesis, namely the weak and the strong thesis. The former is the token identity doctrine. It maintains that every event falling under a mental event also falls under a physical brain event (Kim 1998, p.59). Putting the idea somewhat differently, token identity suggests that any event or occurrence with a mental property also involves some physical brain property (ibid, p. 65). The strong version of mind-brain identity doctrine, viz. type

[96] According to Davidson, the notion of supervenience means that the mental property of a thing cannot be changed without altering physical property. Supervenience suggests that there cannot be two events alike in all physical respects but differing in some mental respect (1970, p.8). Yet this notion does not imply that the mental realm can be reduced to the physical brain. Rather, it accommodates situations in which the same mental phenomenon, state or process can be instantiated (executed physically) by distinct brain properties, states or processes. This makes a case for the phenomenon of multiple realizability (Fodor 1975), which, in turn, does challenge the doctrine that any mental state can be reduced to brain state or property. One example is that individuals with identical brain architectures can experience the same mental phenomenon like pain and feeling of fear. At the same time, different organisms (e.g. different organism (e.g. rats, chimps or humans) with distinct physical brain architectures can experience similar mental events like feeling fear at the perception of an imminent danger. For details, see Kim (2000).

[97] It is important to stress that the mind-brain identity thesis does not mean that mental event kinds and physical brain event kinds are synonyms.

identity thesis holds that any mental state kind is identical to a physical brain kind and therefore the former can be reduced to the latter.

The theoretical model proposed in this chapter endorses a soft interpretation of the mind-brain identity. Its distinctive feature is to hold that one can be a physicalist in relation to the mind. It entertains the idea that one can be a physicalist in relation to the mind without denying at the outset that some mental states and processes causally relevant to manifest behavior can be realized by distinct neural pathways and recognizing that we still know little about how the brain works to suggest that any mental event, state or process can be easily reduced to a particular brain property, state or activity. As a result, our explanatory account may not square well with the reductionist doctrine of eliminative materialism (i.e. vision that all properties and events of the mind can be reduced to a brain event or process). Unlike some eliminativists, who argue that folk psychological categories of the mind (e.g. desires and beliefs) are myths just like phlogiston that ought to be eliminated by mature neuroscience (Churchland 1983), the suggested theoretical model of choice challenges the vision that contemporary neuroscience serves to replace the folk psychological framework and deliberately opted for developing a two-level explanatory account of decision-making behavior (rather than an account centered on a brain level of analysis). .[98] As a result, a two-level account of choice behavior has the merit of addressing explicitly the complex issue of mental causation. The latter refers to the difficulty in understanding how mental properties, states and events relate to physical brain properties and states and whether they are efficient in the causing of physical events (e.g. manifest choice behavior).

5.5 Prospects of explanatory progress

The theoretical model proposed in this chapter employs the theorizing strategy of re-isolation so as provide a theoretical representation of the machinery for human decision-

[98] Many contemporary neuroscientists do not aim to remove the vocabulary of folk psychology from their accounts of cognitive behavior (Damasio 1994, 1999 and 2003, LeDoux 1998 and 2002, Adolphs 2001). Rather, their work reveals the effort to open up black boxes of accounts grounded on folk psychological categories like preferences, desires, expectations and beliefs. This task is accomplished by attempts to uncover the physical brain foundations of individual's capacity to exercise important mental capacities like thinking and decision-making in the real world.

making that offers understanding of how choice behavior happens in the real world and the important roles that emotions may play in it. It aims to develop an account of behavior that contributes to explanatory advancements in economic analysis.

There are two types of explanatory progress (Mäki 2004). One is in terms of the model/theory's capacity to expand its scope. Another is in terms of its capacity to offer improved causal articulation or causal penetration. Explanatory progress as a matter of scope expansion relies on the idea that an account can provide an explanation for an expanding set of explananda in terms of a compact set of explaining items. In this case, a model or theory that exhibits progress as scope expansion (or unification) is able to 'explain much by little' (or by using the same explanatory factors to derive new explananda phenomena). This version of progress is achieved through efforts to isolate those explaining items that account for a wide range of (apparently unrelated) explananda.[99] Explanatory progress as a matter of improved causal articulation involves, in its turn, the task of developing a body of analysis that improves on previous attempts to pick out those explaining items that play major roles in the causal production of the chosen explanandum phenomenon. This vision of progress is also important since a model/theory that succeeds in tracing the chain of events that give rise to the phenomenon under study will locate the place of the explanandum and its major constituents in a complex world of causal dependencies. One important feature of the notion progress as enhanced causal articulation/penetration is that it is concerned with detailing how items of an explanantia set interact to bring about a certain explanandum.

The suggested variant of Simon's account of human decision-making is mainly concerned with the ideal of progress as causal articulation. The two-level account is an attempt to offer a description of the decision machinery that uncovers mental events and properties (physically executed by brain structures and processing activities) with major roles in triggering actual choice behavior. It purports to yield an explanation of behavior that reveals the causal trajectory of economically relevant phenomena, such as intertemporal choice, decision under risk and prosocial choice. In other words, it aims to

[99] For details, see Kitcher (1989) and Mäki (2001).

yield an 'explanation by mechanism', i.e., an account of choice behavior in terms of its underlying causal processes.[100] The suggested model sees a mechanism as a complex system of brain structures (with information processing capabilities) and mental events and processes that produce a non-random (robust) behavior pattern.[101]

The theoretical model developed in this chapter also pursues the goal of progress as scope expansion defined as the theory's capacity to 'explain much by little'. In the next chapters I show that a two-level account is worth developing given that it may contribute to progress as enhanced causal articulation/penetration and scope expansion in the domains of intertemporal choice, decision under risk and prosocial choice. I will show how a two-level model of decision-making behavior has some unifying properties, since it provides an explanatory scheme under which apparently unrelated phenomena in the domains of intertemporal choice, decision under risk and prosocial choice can be subsumed. Note that the suggested two-level account implies some loss in simplicity; its capacity to deal with an increased number of explananda events results from an increase in the size of the explanantia elements. Then, the great challenge posed to the next chapters is to show that the potential explanatory gains associated with a two-level account may outweigh its costs in terms of simplicity.

It is important to stress that the theoretical model proposed in this chapter largely draws on recent ideas from neurosciences and cognitive evolutionary psychology. There is still quite limited knowledge available about the physical brain underpinnings of human behavior. Yet the explanation of decision-making behavior that I suggest is committed to the view that incorporating a brain level of analysis to the body of behavioral economic

[100] An explanation by mechanism is also called mechanistic explanation. I admit that the term 'mechanism' involves some ambiguity. There is no agreement among philosophers of science and practicing scientists on a satisfactory definition of what the mechanism behind a causal explanation consists of. For an interesting discussion of this issue, see Elster (1983), Glennan (1996 and 2002), Salmon (1998) and Machamer, Craver and Darden (2000). Perhaps for the purposes of economic analysis, we can take Guala and Salanti's (2002, p.4, fn. 7) definition of mechanism as a path through the sequence of causal links connecting important explaining items with the phenomenon under study (explanandum).

[101] For a careful study of complex systems, see Simon (1962) and Bechtel and Richardson (1993).

theory may help economists to isolate those processes or mechanisms that play major roles in the causal production of actual choice behavior.[102]

6. Concluding Remarks

In this chapter, I drew on a process-description theorizing approach so as to come up with a theoretical model that describes how components of the machinery for human decision-making operate to bring about choice behavior. The natural point of departure was Simon's theoretical model of boundedly rational choice. Despite my claim that the standard model of bounded rationality has greater explanatory abilities than the mainstream account of choice behavior, I considered some sources of objection to the way in which Simon's heuristic-driven approach deals with the issues of intentionality and causality (section 2).

Based on insights from neurosciences and contemporary cognitive psychology, I provide a variant of Simon's account of rational choice behavior in the real world. It is built on the hypothesis that a genuine explanation of decision-making behavior requires two levels of analysis (section 3). I went on to show that the proposed explanation of decision-making behavior exposes the significant and robust role emotions play in the causal chain of events that bring about patterns of choice behavior (section 4).

The following are among the conclusions that can be suggested. First, despite the loss in simplicity and parsimony associated with an explanation of choice behavior at two levels of analysis, it paves the way for a genuinely explanatory account of decision-making that unearths the chain of physical brain events and mental events that give rise to actual patterns of choice behavior (anomalies included). Second, the inclusion of a physical brain level of description into economic accounts of choice behavior is worthwhile because it might inform economists about processes playing decisive roles in the occurrence of empirical regularities of choice behavior. The third lesson drawn from this chapter is that the suggested two-level account of decision-making offers an ingenious way of addressing the complex issue of causality and as such it contributes to the pursuit

[102] For an interesting argument on this issue, see Camerer, Prelec and Loewenstein (2005).

of explanatory progress as increased causal articulation/penetration of economic analysis. This is because it eventually allows for an improved grasp of the place of the explanandum and its major constituents in the causal structure of the world. The fourth conclusion is that a two-level model of decision-making behavior yields a detailed mechanism description that promotes a clearer understanding of why emotions are not mere sand in the decision machinery but indispensable explanatory items underlying production of choice behavior. The fifth lesson is that the suggested two-level account gives intelligibility to the complex phenomena of judgment and decision making and might contribute to improved predictions of economic behavior. This is partly due to its capacity to detect the conditions under which certain patterns of choice behavior can happen in the world. Of course the burden of the proof is empirical. With this all in mind, the next chapters empirically investigate the explanatory gains of engaging in reformist theorizing strategies (interpreted as de-isolative and re-isolative moves) that characterize behavioral economic analysis in three important domains of human affairs.

CHAPTER 5
EMOTION AND EXPLANATION OF INTERTEMPORAL CHOICE

> To secure a maximum benefit in life, all future events, all future pleasures or pains, should act upon with us the same force as if they were present, allowance being made for uncertainty…But no human mind is constituted in this perfect way: a future feeling is always less influential than a present one
>
> Jevons, *The Theory of Political Economy*, 1870

1. Introduction

Casual observation as well as systematic empirical research informs us that many departments of human affairs - from consuming non-durable goods to mating or child bearing – involve some intertemporal choices. This is partly due to the existence of tradeoffs between an agent's future and current preferences and their estimations of outcomes associated with them. Yet the pervasive phenomenon of choice over time is still in need of a better explanation than that provided by mainstream economic analysis This chapter briefly analyzes some models of choice that have been developed so as to yield improved explanations (predictions) of actual behavior. Its starting point is the view that amended models and theories arise in response to economists' and psychologists' doubts about the sufficiency or even the necessity of some explanantia items (isolated by particular behavioral assumptions) for a satisfactory account of intertemporal choice given its scientific purposes.

The main point of the chapter is to suggest that a behavioral model that specifies how emotion and higher order cognition interact to produce actual intertemporal choice behavior might contribute to the goal of explanatory progress as scope expansion and causal articulation of economic analysis. Additionally, this chapter investigates the potentials of a two-level explanation of choice over time that describes how emotional and higher order cognitive processes and states are related in a complex world of causal

(inter)dependencies. In so doing, it confirms the hypothesis that emotions play significant rather than negligible roles in the (causal) production of decision-making in economically significant domains of human affairs. Furthermore, it demonstrates that a two-level account may lead to a genuinely explanatory theory of choice that attains the goals of scope expansion and increased causal articulation. The remainder of this chapter falls into six sections.

Section 2 presents the standard account of intertemporal choice viz. Samuelson's discounted-utility model (henceforth: DU model) that remains silent about how emotion and cognition bring about actual behavior. It briefly discusses some assumptions that serve to remove some explanatory items that seem to play major roles in the production of economically relevant patterns of behavior. Section 3 scrutinizes some reformed models of choice, interpreted as de-isolative and re-isolative moves, which first arise to deal with recurrent and economically relevant intertemporal anomalies. It claims that amended models of choice increasingly recognize that some explaining items considered by the DU model are not sufficient or even necessary for a genuine account of intertemporal judgments and choices in the real world. It goes on to suggest that advancements in models of intertemporal choice contribute to explanatory progress as scope expansion and increasingly recognize the non-negligible roles emotions may play in some puzzling patterns of behavior. Section 4 analyzes Loewenstein's account of projection bias and the implications of his attempt to describe some affective and cognitive processes underlying anomalies. It argues that his empirically grounded account may better predict (explain) behavior patterns that remain unpredicted and/or unexplained by the DU framework. Inspired by the vision that the existing behavioral models of intertemporal choice do not sufficiently specify how emotions participate in a causal chain of events that produce actual choice over time, section 5 sketches a two-level account that may fill this blank. Section 6 summarizes and concludes the argument.

2. The standard explanation of intertemporal choice

In a paper published by the *Review of Social Studies*, Paul Samuelson provided a formal treatment of intertemporal choice, the so-called discounted utility (DU) model. This

model is built on some idealizing and simplifying assumptions that do not serve to provide a complete description of the complex phenomenon of choice over time. Rather, the DU model's assumptions purport to isolate those explanatory factors with major roles in the production of intertemporal choice and, in so doing, allow for a theoretical representation that yields accurate predictions of economically relevant patterns of behavior.[103]

2.1 The building blocks of the DU model

The DU model is regarded as the standard account of choice over time. Based on a relatively compact set of behavioral assumptions, it hypothesizes that the decision-making agent behaves as if she/he maximized a discounted utility function. It is important to stress that the DU model's assumptions serve to remove some explanatory factors that are thought to play negligible or (causally) unimportant roles in actual intertemporal choice at the aggregate (market) level (Samuelson 1937). The main idealizations and simplifications underlying the DU approach can be summarized by the following set of assumptions: [104]

A1. Integration of new course of action with previously existing consumption plans

The agent does not evaluate prospects in isolation, but based on how a particular new alternative can change his or her pattern of consumption in all periods.

A2. Utility Independence

The agent's overall utility is identical to the discounted sum of the utilities in each period of time. In this case, the agent is thought to exhibit no special preference for patterns of utility across time.

---

[103] As Samuelson put it,

> It is completely arbitrary to assume that an individual behaves so as to maximize an integral of the form envisaged in [DU]. This involves the assumption that every instance in time the individual's satisfaction depends only upon consumption at that time, and that furthermore, the individual tries to maximize the sum of instantaneous satisfactions to some comparable base by time discount (1937, p. 159)

[104] My list of DU assumptions is based on Frederick, Loewenstein and O'Donoghue (2003). I also use some of their notations (pp. 20-24).

A3. Consumption Independence

An individual's utility associated with consumption of a good at a period t + k is independent of her or his consumption in any other period. Therefore, one's comparison between her consumption at period τ and τ' (i.e. her marginal rate of substitution) is completely independent of consumption in τ''.

A4. Discounted utility function is stationary, constant across time

The agent's well-being associated with any consumption or activity is identical in different periods of time.

A5. The agent's time preference discounting is independent from her or his consumption

The discounted utility function is not responsive to different types of consumption. This means that the agent is assumed not to discount utility from different sources and at different rates.

A6. The agent's discounting is constant and her or his intertemporal preferences are consistent

The individual evaluates time in an even-handed fashion, i.e., if in a period t she or he prefers A at period τ to B at period τ + d (for a certain τ and a constant amount d), then in the period t she or he also opts for A at τ to B at τ + d for all possible τ. This implies that individual preferences between outcomes are not changed if outcomes are delayed or anticipated by a common factor.

A7. The individual's marginal utility is decreasing

The utility function is concave, i.e., individuals are thought to spread their patterns of consumption over time.

A8. Individuals have positive time preferences

The agent's discount rate is positive, i.e., it assumes that agent's preferences are biased towards immediate consumption (rather than future and higher gratification).

It is important to stress that the assumptions A1 to A8 play quite specific roles in theorizing and explanation of intertemporal decision-making. Together these assumptions remove some factors from the set of explanatory items (i.e. *explanantia*) that are thought to play minor (if any) role in intertemporal choice. In the next section, I will show that many of the above assumptions will be relaxed once economists and psychologists have doubts about their capacity to isolate explaining factors that are sufficient or even necessary for an adequate representation of actual choice over time. Before that it might be useful to elaborate a bit on what explaining items the DU model seems to include and to exclude. To undertake this task, I pay some attention to the functions of the DU model's assumptions.

Pointing out that an assumption helps exclude a factor is not to yet to justify such exclusion. Other tools are needed for the latter task. To meet this purpose, I am employing the classification of assumptions based on what claims they help make: negligibility assumption, applicability assumption, and early-step assumption (Musgrave 1981, Mäki 2000). More recently, Hindriks (2005) proposes to supplement Musgrave-Mäki's classification of assumptions by including another important type of assumption that is justified by claims about tractability, viz. tractability assumption.

A negligibility assumption justifies the exclusion of a certain explanatory factor (F) based on the belief that its effect on the phenomenon under study is negligibly small given the model's purposes. In this case, claims about negligibility can be true or false and it is an empirical matter to inquire into their truth status. Some behavioral economists often seem to imply that a factor excluded by the conventional model of choice like non-constant discounting has non-negligible impact on behavior, therefore should be included in an improved theoretical representation of choice over time. An applicability assumption justifies the exclusion of a certain explanatory factor based on the idea that a model or theory applies only if the factor is absent. This type of assumption differs from claims about negligibility since the excluded item is thought to play significant (non-negligible) roles in production of the phenomenon under study. Yet the application of the model/theory is possible only in cases where the factor is not present, or is present but

negligibly weak. Applicability assumptions can be either true or false but behavioral researchers, for the sake of empirical testability, work hard to come up with assumptions that are true as often as possible (Mäki 2000, p. 325). An early step assumption justifies the exclusion of a factor based on the vision that, perhaps for analytical convenience, one first removes a certain explanatory item F and will incorporate it into the previously selected set of explainers once a model or theory is further developed. Finally, a tractability assumption is imposed for reasons of tractability, i.e. it is imposed to circumvent some tractability problems associated with formal derivation of a well-behaved objective function, determination of boundary conditions, and so forth. In the next lines I show that, most of the time, economists (behavioral economists included) impose some psychological assumptions due to reasons of negligibility, tractability and applicability. I go on to suggest that behavioral economists and economic psychologists are very aware that relaxation of some DU assumptions, often justified in terms of tractability and neligibility claims might be conducive to the attainment of progress (Frederik, Loewenstein and O'Donoghue 2003, and Loewenstein and Angner 2003).

I take the integration assumption (A1) as an example of applicability assumption. This means that the DU model is only applicable to situations in which individuals can integrate new options with his or her existing consumption plans (i.e. when integration holds). It may be with this in mind Frederick, Loewenstein and O'Donoghue emphasize that integration is a central assumption in most models of intertemporal assumption; it guarantees the possibility of judgments and decisions over time (2003, p.20).

Utility independence (A2) can be regarded as a negligibility assumption. It seems to suggest that the explanatory factor 'the agent's well-being in period t is dependent on her or his past, current and expectations of future well-being' play negligibly small effects on intertemporal choice given the DU model's cognitive purposes and for that reason it can be excluded. This assumption seems to be justified by the belief that only discounting shapes distribution of utility across time in a significant fashion. With this in mind, economists can entertain the idea that the overall utility of a sequence of outcomes is identical to the discounted sum of utility. This assumption is not very discussed by

economists but this may be because it reflects economists' concern with removing any type of preference for patterns of utility over time because this might bring some tractability problems. If this is so, the imposition of A2 might also be interpreted in terms of tractability reasons (details in Frederik, Loewenstein and O'Donoghue 2003).

Consumption independence (A3) is a quite famous DU idealizing assumption. It is also justified by the idea that the explanatory factor 'individual's comparison between consumption of two goods is dependent on her or his past consumption and estimated future consumption' can be ignored since it exerts negligible effects on the occurrence of patterns of intertemporal behavior with economic relevance. It implies that individuals' preferences over certain intertemporal consumption profiles are not affected by the nature of consumption. This is one of the most debated DU assumptions, since it is consistent with the idea that an agent's preference between a Japanese restaurant and a French one tonight is independent on whether she or he ate French yesterday (or expect it to have it at tomorrow's dinner). Many challenged the empirical accuracy of consumption independence. The very father of the DU model recognized the inexistence of a compelling reason for 'consumption independence'. A3 can be interpreted as a false assumption and this is implied by Samuelson's claim that 'the amount of wine I drank yesterday and will drink tomorrow can be expected to have effects on my today's indifference slope between milk and wine (1952, p.674). Yet consumption independence was not relaxed until the 1980s. In my own interpretation, this is largely so because this assumption circumvents the problem of deriving a tractable and well-behaved discounted utility model. If this is so, there are also grounds for regarding A3 as a tractability assumption. Nowadays behavioral economists realize that the claims about negligibility and tractability associated with A3 can be challenged. This idealizing assumption removes an explanatory item that are significant (rather than causally unimportant) for the occurrence of actual behavior.

The stationary instantaneous (discounted) utility (A4) also seems to be an idealizing assumption that involves claims about negligibility and tractability. Its exclusion of the explanatory factor 'individual preferences change over time in predictable and non-

predictable manners' seems to be justified by the idea that its impact on patterns of behavior at the economy level is negligibly small. A4 is also imposed for tractability reasons. By excluding the explanatory factor 'preferences change over time', it allows for a tractable discounted utility function (Frederik et al 2003).

Independent discounting from consumption (A5) is a clear example of tractability assumption. It seems to justify the exclusion of a non-negligible (significant) explanatory item, viz. 'utility discounting is made at different rates and depends on the agent's perception of the nature and type of consumption' by the idea that a tractable model of intertemporal choice is only possible when the agent's discount function has a particular property - it is invariant across all types of consumption. Frederick, Loewenstein and O'Donoghue's (2003, p. 22) suggest that if this assumption does not hold, the very notion of unitary time preferences loses its bite. This might bring tractability problems to economic models of intertemporal choice, since characterization of a time preference becomes dependent on the thing being delayed, such as 'chocolate time preference', 'summer vacation time preference', and so forth. If this is so, tractability seems to be a crucial (pragmatic) consideration for imposing an unrealistic (false) assumption like A5.

Constant discounting (A6) seems to be a negligibility assumption. It excludes the factor 'individuals have conflicting preferences' from the explanatory set based on the belief that its effect on behavior patterns within the economy is negligibly small. It is also likely that tractability considerations also serve as reasons for imposing A6. Frederik et al emphasize that if the assumption constant discounting is not imposed, formal characterization of an individual's time preferences will become less tractable since it may require specification of an entire discount function and individual time preferences cannot be taken as a single discount rate (Frederik, Loewenstein and O'Donoghue 2003, p. 23). Based on experimental research, behavioral economists challenges the claims about negligibility and tractability associated with constant discounting. Some of them even seem to take A6 as an early step that might be relaxed on behalf of a generalized hyperbolic discount function (Laibson 1994, Laibson, Repetto and Tobacman 1998).

Diminishing marginal utility (A7) is an idealizing assumption that involves claims about negligibility, justified by the idea that the factor 'the individual prefers to concentrate rather than to spread patterns of intertemporal consumption' can be excluded from the analysis since its effect on the phenomenon under study is negligibly small. Economists offer some evidence that support the claim that most people within the economy opt for spreading their consumption over time.

Finally, the assumption of positive time preferences (A8) is an idealization that refers to claims about negligibility. It is justified by the idea that the factor 'the individual prefers future to immediate consumption' plays negligible roles in the production of economically relevant patterns of behavior (and therefore can be removed). More recently, behavioral economists challenge this assumption and take them sometimes as early step assumption that might be relaxed to allow for a model of intertemporal choice that pursues the goal of scope expansion (for details see Frederik et al 2003).

Based on the aforementioned assumptions, the DU model gives an account of choice over time in terms of maximization of a well-behaved discounted utility function that represents an individual's preferences over a certain intertemporal consumption profile.[105] The DU model offers a framework to deal with various instances of consumption (and savings behavior). For instance, the model explains why people live off their own (family) endowments in their early study and working years, why they save more during mid life (their most productive working phase) and why they live off saved income when they get retired (Modigliani and Brumberg 1954). Table 1 offers a summary list of the main DU assumptions and the major explanatory items that they exclude:

[105] The discounted utility function can be written as $U_t(c_1, c_2 \ldots c_T) = \sum_{k=0}^{T-t} D(k).\ u(c_{t+k})$, where $D(k) = (1/1 + \rho)^k$
The function U (.) denotes an additive instantaneous function, the discount function D (k) represents the importance an individual assigns, in period t, to her well-being (consumption) at the period t+k, and u (.) corresponds to an individual's unit of preference satisfaction at a particular period of time.

Main DU assumptions	Excluded explanatory items
1. Integration of new alternatives with existing plans	People do not make well formed plans about streams of future consumption and cannot recalculate the optimal consumption plan whenever she makes an intertemporal choice
2. Utility Independence	The well-being experienced in t is dependent on the agent's well-being in the past and her or his expectations of future well-being'
3. Consumption Independence	Individual comparison between consumption of two goods is dependent on her or his past consumption and estimated future consumption
4. Stationary instantaneous utility	Individual preferences change over time
5. Discounting is independent from consumption	Preferences are discounted at different rates and this is dependent on the nature and type of consumption
6. Constant discounting and time consistency	Individuals have conflicting preferences over time
7. Decreasing marginal utility	Preferences over concentrating (rather than spreading) patterns of intertemporal consumption
8. Positive time preferences	Preferences are biased towards the future

Table 1 – The main assumptions underlying the DU model and the explanatory items that they seem to exclude from the set of explainer

Based on the above attempt to classify the roles of the main DU assumptions, I am led to suggest that tractability considerations play an important influence on certain DU assumptions and even seem to justify the removal of some explanatory factors that do not play negligible effects on the phenomenon under study (for an interesting analysis of the importance of tractability considerations, see Hindriks 2005).

3. Empirical evidence against the discounted utility model

In this section, I briefly present some empirical results that raise doubts about whether some DU assumptions removed some explanatory items from the set of explainers of intertemporal judgment and decision-making that play non-negligible roles in production of actual behavior patterns.

3.1 The phenomenon of hyperbolic discounting

Behavioral scientists have designed experiments that confirm the hypothesis that individuals often prefer current to future consumption. In contrast to predictions made by the DU model, actual individuals discount utility at decreasing rates when their time horizon is ample (Ainslie and Herrnstein 1981, Thaler 1981, Elster and Loewenstein 1992). These findings indicate that actual individuals seem to exhibit a hyperbolic discount function (i.e. discount rates lower as time goes by).[106] In response to systematic and recurrent evidence that the DU model failed to deal with important patterns of behavior involving preference reversals, economists and psychologists began to doubt whether the assumption of constant discounting and time consistency (A6) excluded an item with major (instead of minor) causal roles in the production of intertemporal decision-making given the model's pursue cognitive tasks (Ainslie 1992, Laibson 1997, Frederick et al 2003).

3.2. Small amounts of money are discounted at high rates vis-à-vis large sums

There are patterns of intertemporal choice behavior that failed to be explained by the standard DU model. Thaler (1981) found evidence that individuals are often indifferent to receiving $15 now or $60 in a year, $250 immediately or $350 in a year, and $3000 now or $4000 in a year. These empirical results reveal that the ways in which actual individuals discount future utilities (outcomes) are dependent on the amount of money at stake. Note that the discount rate associated with the choice between $15 and $60 is 139%, whereas the discount rate associated with $3000 and $4000 is approximately

106 For example, Thaler (1981) conducted various experiments showing that individuals often prefer an apple today to two apples tomorrow, or prefer two apples in 51 days to one apple in 50 days. In one of his experiments, participants are asked how much money they would be willing to receive in 30 days, 1 year or 10 years to become indifferent to gaining $15 immediately. The answers suggest that an individual may rely upon a hyperbole-shaped discount function to tradeoff present and future preferences. The average amount requested for the month delay was $20; one year of delay $50, and $200 for a 10 year-horizon. These findings reveal average discount rates of 345% (month-interval), 120% (delay of a year) and 19% (within a 10 year time-horizon). Thaler and his collaborators also asked participants if they would prefer to be given a $100 check to be cashed immediately or a $200 check to be cashed in two years. Most preferred immediate consumption in this setting but preferred a $200 check that could not be cashed for 8 years to a $100 check to be cashed in 6 years. These behavior patterns deviate from DU model predictions. The above puzzling phenomena are known as instances of the common difference effect. For other examples of behavior patterns revealing hyperbolic discounting, see Benzion, Rapoport and Yagil (1989).

29%.[107] Such puzzles to the discounted utility model are called 'magnitude effect'. [108] They motivate decision theorists to consider whether the removed explaining items by the utility independence (A2), consumption independence (A3) and constant discounting (A6) assumptions play non-negligible effects on economically relevant patterns of intertemporal behavior and therefore ought to be taken seriously by an explanatory model of choice.

3.3 Individuals discount gains more highly than losses

Experiments informed by economics and psychology discovered an asymmetric relationship between discounted losses and discounted gains. Thaler (1981) shows that individuals discount gains at rates from 3 to 10 times higher than those associated with losses. Along similar lines, Loewenstein (1988) designed an experiment in which participants were found to be indifferent about receiving $10 now or $21 in a year, and at the same time remained indifferent to losing $10 immediately or $15 in a year. Subjects also remained indifferent to gaining $100 now or $157 in a year, and losing $100 immediately or $133 within a year. The above findings pose enigmas to the discounted utility model called sign effects. Such intertemporal choice anomalies motivate doubts about whether the assumptions of independence (A2, A3) and constant discounting (A5) removed items from the set of explainers that are significant for the occurrence of actual choice over time.

3.4 Patterns of intertemporal choice that reveal farsighted (rather than myopic) behavior

The DU model allows for predictions that the agent will always prefer immediate to later (higher) gratification (consumption of a good). Loewenstein (1987) found experimental evidence that under many circumstances individuals may prefer to postpone a rewarding consumption and even to anticipate a negative outcome. Such puzzling patterns of behavior are called the savoring effect and the dread effect, respectively. This first case of DU anomaly shows that individuals derive utility from expectation of a future rewarding

107 Very similar empirical results are also encountered in Ainslie and Haendel (1983), Loewenstein (1987), Raineri and Hachlin (1993) among others.
108 For more evidence of the magnitude effect, see Holcomb and Nelson (1989) and also Benzion, Rapoport and Yagil (1989).

outcome, whereas the other indicates that agents experience negative utilities from the expectation of consumption of a punishing outcome.[109] Both patterns of behavior throw light on the possibility of individual preferences biased towards the future (a case for farsightedness).[110] When recurrent and systematic, these empirical anomalies prompt economists (among other behavioral researchers) to doubt whether the assumption of positive time preferences (A8) did remove an item from the explanatory set with non-negligible effects on the phenomenon of choice over time in the real world.

3.5 The asymmetric relationship between willingness to postpone or bring forward consumption

The DU model is built on the idea that the ways in which utilities are distributed over time do not matter to decision-makers, since a sequence of outcomes is identical to the sum of discounted utilities. This leads to the prediction that the price the agent is willing to pay to obtain a good immediately (i.e. in-advance cost) is the same as the delay premium (i.e. the price the agent is willing to pay to postpone gratification). However, empirical studies made by Loewenstein (1988, pp. 200-201) show cases in which the agent's willingness to speed up consumption differs from that to delay gratification. In his experiments, subjects are asked how much they would pay to have a VCR immediately, delay delivery of a VCR (already bought) for one year, or own a VCR in one year (ibid, 205-6). Loewenstein's findings question the discounted utility model since the compensation the agents required for a delay with receiving a good they had already bought was higher than the amount they were willing to pay for earlier consumption of the VCR (i.e. the difference between the price of the good for immediate consumption

109 One example of savoring effect is keeping a bottle of wine for a special occasion in the future. Dread effect, in its turn, occurs when individual choices reveal agent's preference for anticipating a future loss (or punishing outcome) rather than delaying it. The phenomenon of saving after retirement is an economics-related illustration of dread effect (see Loewenstein 1996).

110 In order to test the hypothesis of negative time preferences, Loewenstein (1987) asked undergraduate students whether they would prefer to be kissed by their favorite movie-star immediately or with some delay. Subjects were also asked if they would prefer to receive a non-lethal 110 volt shock immediately or later. The experimenters found that participants preferred to postpone the pleasant situation (i.e. being kissed by their favorite actress or actor) rather than anticipate it. Individuals also preferred anticipating an unpleasant experience (i.e. getting an electric shock) to postponing it.

and the price for delayed gratification).[111] These results lead to questions about whether the assumptions of independence (A2 and A3), stationary instantaneous utility function (A4) and time consistency (A5) excluded explanatory items with significant roles for actually observed choice over time and therefore they ought to be supplemented or even replaced by new explanatory elements.

3.6 The impact of previous and current consumption

The standard explanation of intertemporal choice presupposes that individual's future preferences for a particular good or service are independent of her or his previous and current consumption. [112] Furthermore, it draws on the idea that the overall value (i.e. experienced well-being) that an individual ascribes to a sequence of outcomes is identical to the discounted sum of utilities in each period involved. In an experiment Loewenstein and Prelec (1993) asked subjects to choose over spending their evenings for 5 weekends in sequences of dinners (e.g. eat at home, at a fancy French restaurant, at a fancy lobster restaurant). They found that when facing with a intertemporal profile of the type I "Fancy French restaurant in the first; Eat at home in the second, Eat at home third; Eat at home in the fourth and Eat at home in the fifth" and with another of type II "Eat at home in the first, Eat at home in the second, Fancy French Restaurant in the third, Eat at home in the fourth and Eat at home in the fifth", most people (89% of participants) prefer the second option (II). Yet when individuals have to choose between a prospect type III "Fancy French restaurant in the first, Eat at home in the second, Eat at home in the third, Eat at home in the fourth and Eat in a Fancy Lobster Restaurant in the fifth" and another IV "Eat at home in the first, Eat at home in the second, Eat in a Fancy French Restaurant in the third, Eat at home in the fourth, Eat at a Fancy Lobster Restaurant in the fifth", they can exhibit patterns of behavior regarded as anomalies. The DU model predicts that an individual who prefers profile II ought to select IV. Nevertheless, Loewenstein and Prelec

111 Based on three different experiments studying frame of mind, Loewenstein (1988) discovered that subjects often required an amount to compensate for delaying consumption of a good purchased (e.g. video cassette recorder or restaurant voucher) from 2 to 4 times greater than the amount they would sacrifice to speed up consumption.

112 This statement serves simplifying purposes. Even Samuelson (1952) recognized that independence assumptions may offer a poor description of individual's structure of time preferences. According to him, "the amount of wine I drank yesterday and will drink tomorrow can be expected to have effects on my today's indifference slope between wine and milk (p.674).

found that many people (49%) prefer consumption profile II and profile III (instead of IV). This empirical result motivates economists and psychologists to investigate further whether the assumptions of independence (A2 and A3) removed explanatory factors with non-negligible (causal) roles in choice over time.[113]

3.7 Time-conflicting preferences

The DU model presupposes that individuals exhibit patterns of behavior as if their preferences were intertemporally consistent.[114] This simplifying idea helps one derive a 'well-behaved' discounted utility function. Nevertheless, there is evidence that many patterns of behavior failed to be predicted by the DU model. For instance, Christensen-Szalanski (1984) designed an experiment in which pregnant women are asked whether they would like to have nitrous oxide (anesthesia) during childbirth. Subjects thought that they would not, but with the onset of labor, their preferences shifted and they asked to be anesthetized. According to the experimenters, this experimental result reveal preference reversals in intertemporal choice.[115] Nowadays behavioral economists have collected economically important patterns of consumption and savings behavior that cannot be predicted by the discounted utility framework. For instance, actual individuals have high credit card debts and at the same time allocate a huge portion of their income to illiquid assets (e.g. retirement accounts, housing schemes, Christmas clubs).[116] Such patterns of behavior come to question whether some DU assumptions like constant discounting excluded an explanatory item (time-conflicting preferences) with major (causal) roles in actually observed patterns of behavior.[117]

---

[113] For details, see Loewenstein and Prelec (1993).
[114] The model is built on the idealizing assumption that individual discount future utilities at a constant rate (i.e. based on an exponential discount function).
[115] Perhaps a skeptical economist may not take this experiment as significant evidence against the predictive power of the discounted utility model because it does not refer to an economically relevant phenomenon.
116For details see Loewenstein and Elster (1992) and Angeletos et al (2001/ 2003).
117 Elster and Loewenstein (1992) offer evidence that individuals are sometimes aware of their time conflicting preferences and therefore rely upon commitment strategies that lead them to resist temptation of immediate gratification and to opt for delayed (higher) reward. Some like Schelling (1984), Elster (1985) and Ainslie (1992) examine the possibility of explaining intertemporally inconsistent behavior by reference to 'multiple selves'.

Table 2 summarizes systematic patterns of intertemporal choice behavior that are regarded as anomalies (they cannot be accommodated by the standard DU model) and offers an interpretation of why they might happen.

Some patterns of behavior that cannot be accommodated by the DU	**Interpretation of why this happens**
People often prefer immediate to delayed consumption, but the rate at which agents discount future utilities lower with the time horizon (i.e. Common difference effect) People overborrow in the revolving credit market and at the same time allocate part of income in illiquid assets	The assumption of constant (exponential) discounting is an excessive simplification of individual time discounting
Small amounts of money are discounted more highly than large sums (magnitude effect) Individuals discount gains more highly than losses (sign effect)	The independence and constant discounting assumptions allow for derivation of a model of choice with an oversimplified representation of the structure of individual time preferences. It ignores the important role played by the individual's perception of losses or gains (and their amounts) in choice over time.
Intertemporal choice of behavior reveals savoring and dread effects Individuals prefer keeping a bottle of wine for a special occasion. Agents go on saving even after retirement	The DU model downplayed the significance of situations in which individuals reveal negative time preferences (farsightedness)
The price that the agent is willing to accept in order to postpone consumption of a good that she or he has is different from the price she or he is willing to pay for speeding up consumption of a good that she or he does not have	The standard model neglected the important role of individual's perception of loss associated with postponing a good owned and of gain associated with anticipation of consumption of a good she/he does not possess.
Individual preferences for a particular intertemporal prospect are influenced by her or his past and future consumption.	The DU model failed to offer a correct depiction of the explanatory factors laying major roles in shaping individual's time preference and discounting since it downplayed the importance of earlier consumption on individual's preference for future outcomes
Agents prefer to save money for old age but when they see some consumer goods that elicit strong feelings they opt for immediate gratification. People prefer to put their money in Christmas clubs and compulsory retirement funds to commit themselves to higher delayed consumption.	People prefer to put their money in Christmas clubs and compulsory retirement funds to commit themselves to higher delayed consumption.

Table 2- List of puzzling instances of intertemporal choice behavior with interpretations of why they happen

Given that the above deviations from the DU model are recurrent and systematic, they led economists to consider that some assumptions of the DU model ought to be relaxed since it removed explanatory elements with major roles in production of actual choice over time.

4. Behavioral models that come to grips with DU anomalies

In this section I briefly discuss a sequence of models of choice that can be intepreted as instances of de-isolation that serve to relax some DU assumptions that failed to consider those key explanatory factors to the occurrence of intertemporal judgment and decision-making and eventually contribute to improved predictions and explanations of manifest choice behavior.

4.1 The amended DU model with hyperbolic discounting

The hyperbolic discounting model relaxes the assumption of constant discounting (A6) so as to come up with an account of choice over time that deals with behaviors covered by the standard DU model and also those regarded as anomalies, such as preference reversals. This reformed version of the DU model includes a previously excluded item 'individuals have conflicting time preferences' that seems to play important roles in production of behavior patterns regarded as anomalies. Unlike the standard DU model that represents discounting by a constant rate (exponential discount function) due to analytical convenience, this variant of the DU model assumes that the relative weight that the agent ascribes in period t to her or his well-being in period t + k can be expressed by a hyperbolic discount function (which can accommodate exponential discounting as a particular case).

The novelty of model of hyperbolic discounting is that the inclusion of the assumption of hyperbolic discount function takes into account explanatory items "individuals have time conflicting preferences' and that 'an individual prefers a delayed (larger) reward when choice involves a long delay but opts for immediate (lower) gratification when the future becomes imminent'. In this case, this amended model of choice supplements the set of DU assumptions (A1-A8) with the assumption 'individuals discount future utility

hyperbolically)' so as to incorporate an explanatory item that help us to deal explain important patterns of choice behavior.

For instance, the hyperbolic discounting model predicts that pregnant women who reject anesthesia will change their minds as soon as labor starts. The model implies that people prefer a delayed (bigger) reward when the time delay involved is long. When time delay approaches zero, their preferences become biased toward immediate (though lower) gratification.[118] The model of hyperbolic discounting yields satisfactory predictions of economic behavior, and the empirical evidence is quite supportive. The generalized model of hyperbolic discounting predicts that American individuals will consume more goods than they expect to do. David Laibson and his colleagues argue that in a world without illiquid assets, a hyperbolic discounting model resolves puzzling consumption-saving behaviors such as co-movement of income and consumption, high rates of credit card debts in revolving credit markets, the phenomenon of asset-specific propensity to consume, low levels of precautionary savings, the strong correlation of patience levels with income, wealth and even age (Laibson 1994, Angeletos et al 2001). The hyperbolic discounting model accommodates phenomena such as procrastination and overconsumption of addictive goods (O'Donoghue and Rabin 1999 and 2000). In addition it predicts new phenomena like strategic ignorance (i.e. why people sometimes do not acquire information even when it is free).[119] This puzzling phenomenon can be explained in terms of the idea that agents, who discount future consumption "hyperbolically", may be concerned with withdrawing from a satisfactory course of action when its costs become imminent and therefore opts no as this serves the role of a "commitment strategy".

As we can see, the hyperbolic discounting model can deal with explanandum phenomena that were viewed as DU anomalies (e.g. preference reversals, why people overborrow in the revolving credit card market and at the same time buy illiquid assets and try to

118 Studies with drug addicts and mentally handicapped teenagers corroborate the empirical hypothesis of hyperbolic discounting. For details about the early experiments and their interpretations, see Ainslie and Haendel (1983).
119 For details, see Carrillo and Mariotti (2000).

commit themselves to investments in Christmas clubs) because it incorporates a previously excluded explanatory item "conflicting time preferences' significant for the occurrence of some choice behavior. The DU model became very popular among economists among other social scientists due to its capacity to offer an answer to two important questions. The first concerns the insufficiency of the explanatory items isolated by main DU assumptions for an account of a set of phenomenon that the DU model aim to explain (but sometimes fail to do so). Another has to do with the question whether the explaining items isolated by the DU assumptions are sufficient to account for new explananda events. The experimental literature on choice over time indicates that the generalized DU model with hyperbolic discounting offers improved predictions of actual behavior and may even allow for progress as scope expansion.[120, 121]

4.2 The reference-point model

Inspired by Kahneman and Tversky's prospect theory, Loewenstein and Prelec (1992) came up with a model that aims to yield an empirically valid representation of choice over time. The reference point model of intertemporal choice draws on a reformist theorizing strategy that doubts whether one or more items included within the DU model are unnecessary to warrant a satisfactory account of manifest choice behavior and therefore ought to be replaced by previously excluded explanatory factors. This reformed model relaxes the assumptions of utility and consumption independence (A2 and A3) and stationary (well-behaved) instantaneous utility function (A4). Alternatively it assumes that 'individuals are equipped with bounded rationality and therefore they perceive intertemporal prospects as losses or gains, viz. deviations from actual consumption in the period τ from the reference point'. In addition, the reference-point model presupposes that individuals make decisions based on an s-shaped value function that is concave for perceived gains and convex for losses (instead of a well-behaved instantaneous utility

120 According to Frank (1992), discounting at hyperbolic rates helps explain adaptive behavior responses. In an environment of extreme uncertainty in relation to the future, a bird in the hand may be worth two in the bush.

121 Yet there are some sources of criticism to the DU model. It offers an explanation in terms of parameters of time delay and reward profiles and as such remains silent about how exactly intertemporally inconsistent preferences and choice are generated. The model does not offer understanding of why certain types of consumption objects often produce time-inconsistent behavior while others do not. For details, see Loewenstein (1988, pp. 21-2).

function). These newly included assumptions purport to replace some previously included explanatory items like 'the agent's well-being in period t is independent on past utilities and estimated well-being in the future', 'the agent's comparison between two goods is independent on her past and future consumption' and 'individual preferences do not change over time' with previously excluded explanatory factors with major roles in actual behavior, such as 'one's currently experienced well-being depends on her past and estimated well-being', 'individual judgments of intertemporal prospects are influenced by the agent's past and expected (future) consumption, and 'choices are made based on preferences that change over time'.

The reference-point model accommodates phenomena regarded as DU anomalies, such as sign effect, magnitude effect, common difference and time-preference reversals (Loewenstein and Thaler 1989).This reformed account of the intertemporal choice model enables us to examine a wide set of explananda events (Loewenstein and Prelec 1992).[122, 123]

4.3 The anticipatory utility model

This empirically grounded account of intertemporal choice is built on a strategy of de-isolation. Its aim is to accommodate economically important DU anomalies and therefore to enhance the empirical validity of economic analysis of intertemporal choice. The model purports to be more concrete than the DU model, since it relaxes some assumptions like positive time preferences (A8), independence assumptions (A2, A3), stationary utility function (A4) and constant discounting (A6) that consider some explanatory items that turn out to be insufficient to deal with economically important patterns of behavior. Relaxation of the above DU assumptions serve to supplement the explanatory set of the DU model with previously excluded items playing non-negligible effects on actual behavior (e.g. dependence, dynamic inconsistency of preferences and

122 Since the reference-point model relaxes the assumptions of independence and constant discounting, it can predict instances of preference reversals that the hyperbolic discounting model accommodates. Yet most economists prefer to use the amended DU model with a general (hyperbolic) discount function.
123 Loewenstein (1988) acknowledges that the reference-point model resembles the DU approach in some aspects. For instance, it also assumes that individual preferences over intertemporal prospects are separable. The reference-point model also remains silent about why and how certain goods and types of consumption evoke more impulsive or myopic behaviors than others.

preferences that shift over time). According to Loewenstein (1987), this alternative model of choice improves the way in which the discounted utility approach represents theoretically agent's time preferences and discounting (Loewenstein 1987).

The anticipatory utility model can deal with phenomena that the DU model cannot do, such as why people prefer to delay consumption of certain rewarding goods and anticipate particular undesirable outcomes. The model can predict DU anomalies like savoring and dread effects. In addition, the anticipatory utility model can explain why people who planned to consume a good after some delay will go on saving it even after the time for consumption arrives (ibid, p.39). Furthermore, it offers understanding of why individuals discount future consumption of different goods at different rates.[124] The anticipatory utility model offers a different account of magnitude effect. It explains this DU anomaly by reference to the idea that people discount small amounts of money more steeply because of the dread reactions associated with large amounts. The model also offers a similar interpretation of the sign effect: it suggests that it will happen when dreading future negative consequences is a stronger emotional response than savoring positive future outcomes (ibid.).

The anticipatory utility account can deal with an economically important anomaly called 'savings behavior after retirement'. In Loewenstein's (2000) own words,

> Retirement for the young is a non-vivid event – perhaps partly because thinking about old age is aversive and tends to be avoided… As retirement approaches, however, the prospect of having inadequate funds for retirement becomes increasingly vivid and causes anxiety, anxiety that can be allayed in part by stepping up savings. The onset of retirement itself, and the sudden loss of wage income, of course, greatly increases this anxiety for the future. This anxiety raises the return in terms of anxiety reduction, of saving, and counteracts the savings-discouraging effects of the loss of income upon retirement. (pp. 676-7)

The above passage schematically offers an account of this deviation from the DU model in terms of the dread feeling reactions. Individual's anticipation of negative utilities associated with lower incomes that follow retirement explains why savings after

124 Acording to Frederick, Loewenstein and O'Donoghue (2003), people discount different goods at different rates "because utility from anticipation creates a downward bias on estimated discounted rates and this downward bias is larger for goods that create more anticipatory utility" (p. 38).

retirement happens. Thus there are grounds for suggesting that the anticipatory utility model questions the sufficiency of the isolated set of explanantia that characterizes the DU model and incorporates new ideas that allows explaining events that the DU model failed to cover and also to accommodate new phenomena. This being the case, the anticipatory utility model may contribute to predictive and eventually explanatory improvements in economic accounts of intertemporal choice.

4.4 A visceral model of intertemporal choice

Loewenstein (1996) provides a behavioral model built on a de-isolation strategy that relaxes some DU assumptions, such as utility and consumption independence (A2, A3), stationary utility function (A4), independent discounting from consumption (A5) and constant discounting (A6). This behavioral model of intertemporal choice purports to supplement the set of explainers of the DU framework with some previously excluded explanatory factors that play major roles in production of actual behavior. Some examples of items excluded by some relaxed DU assumption are (a) individual's wellbeing is shaped by her or his past experiences and predicted future utilities and feelings; (b) her comparisons between alternative courses of action is contingent on her past and future consumption; and (c) she discounts future at different rates depending on the nature of the consumption or good and on different visceral states attached to them).

The visceral model hypothesizes that individual discounted utility is state-dependent, i.e., shaped by the agent's currently experienced and estimated visceral states (pp.274-278). The model relies on two main premises. One is that immediate visceral factors have a disproportionate effect on behavior and therefore sometimes crowd-out individual's goals other than those directly associated with the visceral state. Another is that individuals underestimate the strength and duration of their own and others' earlier hedonic experiences and often mispredict their future preferences and feelings. The model assumes that the desirability of a reward as being dependent on time delay as well as the situational factors involved. The merit of these new assumptions is that they include explanatory items like sources of utility and consumption dependency, various sources of

time inconsistence (conflicting time preferences) and possibility of discounting based on different sources and at distinct rates.[125]

The visceral model arise in response to questions about the sufficiency of the isolated set of explanantia items of the DU model for an account of actual intertemporal choice. The model offers rather intuitive explanations of economically significant DU anomalies. For instance, it accounts for the phenomenon of time conflicting preferences by referring to the idea that real-world people have difficulty in effectively predicting their future utilities and payoffs when they experience a strong visceral state or affective experience. Unlike other variants of the DU model, the visceral model offers an account of why (and under what conditions) some individuals exhibit myopic or farsighted choice behaviors. It suggests that certain goods elicit strong emotions and affective states that may alter an agent's perceived tradeoff between current and future outcomes and eventually lead to a bias towards immediate consumption, whereas other goods elicit visceral factors associated with dread and savoring states conducive to farsightedness.

The visceral model accommodates instances of choice behavior resulting from self-control and impulsivity problems. For example, it yields an explanatory account of time preference reversals, such as those found in Christensen-Szalanski's experiment with pregnant women. The visceral model offers understanding of why and how time conflicting preferences and intertemporally inconsistent patterns come about. It suggests that visceral factors such as pain and fear in labor motivate an expectant mother to opt for anesthesia. In other words, the mother's experience of negative visceral states alters her perception of the relative desirability of avoiding immediate pain during childbirth (i.e. taking anesthesia).

The visceral model offers an explanatory account of why reward-delaying behavior happens. It predicts that the existence of certain cues signaling one's proximity to a desired (rewarding) object, such as physical proximity, will activate strong visceral states

125 See Loewenstein 1996, p.289.

(e.g. appetites and cravings) that will bias the agent's preferences for immediate gratification. [126, 127]

In this case, a visceral model of intertemporal choice offers an improved account of actual behavior because it includes explanatory items that play major roles in intertemporal judgment and decision-making like 'types of dependence', 'bodily-driven sources of time inconsistency, 'changing preferences over time in predictable and non-predictable manners', 'discounting that is dependent on the nature of the good and type of consumption', etc.[128]

4.5 The model of projection bias

Recently, Loewenstein, O'Donoghue and Rabin sketched a behavioral model that aims to uncover mental processes and states with significant roles for the occurrence of actual behavior. This model arises in response to questions about the necessity of the isolated explaining items excluded by the assumptions of independence (A2 and A3), stationary instantaneous utility (A4), discounting independent from consumption (A5), constant discounting (A6), and positive time preferences (A8).

The account of intertemporal choice in terms of the projection bias seems to replace the principle of discounted utility maximization by an empirical hypothesis: an individual relies on her or his current (visceral) states to predict their future preferences and the (hedonic) utility outcomes associated with available courses of action. The novelty of the model of projection bias is that it replaces the axiomatic version of discounted utility approach by a heuristic type of account. The model relies on the behavioral assumption that individuals make intertemporal choices using cue-based heuristics of the type

126 Mischel and his associates (Mischel 1974, Mischel, Shoda and Yuich 1992) asked children to reveal their favorite food and were told that they would receive what they wanted if they waited 15 minutes for a confederate. Otherwise, they would receive another food anytime.

127 They also found that kids entertaining themselves with fun toys and other distractions experienced visceral states that shaped their choice over a delayed higher reward.

128 Yet there may be some sources of objection to the explanatory capabilities of the state-dependent utility model. Like other variants of the DU model explanatory factors included in the body of theory visceral model are arguments of the individual instantaneous utility function. Perhaps a genuinely explanatory account of choice over time requires an explicit description of cognitive and affective processes involved in the causal chain of events that produce actual intertemporal choice.

"project current states onto future feelings, preferences, and goals". This theoretical statement may be viewed as a matter of including non-negligible explanatory factors previously excluded by the conventional model of intertemporal choice. The model of intertemporal choice with projection bias can be regarded as a development of Loewenstein's visceral (state-dependent) utility model.

The model of projection bias assumes that an individual's objective function can be represented by u (c_τ, z_τ), where z_τ is the state variable that captures the hedonic effects of past consumption.[129] The model suggests that a projection bias will happen whenever the agent uses her or his current visceral state z_t to estimate future utility.[130] It assumes that the agent's at period t experiencing a visceral state denoted z_t will predict that her or his utility in period τ given a certain visceral state z_τ – u^p (c_τ; z_τ/ z_t) - will be between her or his true future utility - u (c_τ, z_τ) - and her or his utility attached to the currently experienced state – u (c_τ, z_t) (Loewenstein, O'Donoghue and Rabin 2000, pp. 9-19).[131]

The projection model bias explains the occurrence of intertemporally inconsistent choices by referring to the agent's difficulty in estimating the intensity, duration and impact of current visceral factors (e.g. emotions, feelings, pains, hunger, and cravings) on her/his own and others' future preferences, feelings and utilities (ibid, pp.1-9). It also predicts that an individual under the experience of a 'hot' visceral state (e.g. hunger, sexual arousal and fear) will mis-predict what she (or another) will think, prefer, feel or do under the experience of affectively neutral situations (i.e. those that elicit 'cold' states or cool-blooded deliberations). Even though Loewenstein and his colleagues remain silent about the kind of account that their model implies, my own interpretation is that this analytical model resembles a heuristic approach. The model of projection bias can accommodate

129 This sheds light on the relaxation of the assumptions of independence, stationarity and independence of utility from consumption.
130 This requires us to question the necessity of the explanatory item 'preferences are intertemporally consistent' and to come up with alternative behavioral assumptions that include the factor 'there are dynamically inconsistent preferences'.
131 Putting the issue a bit more formally, the model suggests that up (c□; z□/ zt) = (1-□). u (c□, z□) + □̃u(c□, zt), for all 0 ≤ □ ≤ 1.

many DU anomalies that shed light on the standard model's failure to give a satisfactory theoretical account of the driving forces behind dynamically inconsistent behavior, such as simultaneous over borrowing in credit card markets in the USA with large savings after retirement.

The model of projection bias describes specialized processes that explain how intertemporally inconsistent behavior happens. According to Loewenstein and Angner (2003), there are three sets of cognitive and affect-driven processes that are significant for the observed gap between one's estimated and actual (future) preferences. They are called focalism; the misconstrual problem; under-appreciation of one's ability to adapt preferences to favorable and unfavorable situations and ownership; and hot and cold empathy gap.

Focalism refers to the process by which an individual ascribes high values to salient aspects of a task at hand (Loewenstein and Schkade 1999, p. 94-96 and Loewenstein and Angner 2003, pp. 370-1). It is a source of time inconsistent behavior to the extent that individuals often rely more heavily on a single (or a few) affectively salient element(s) to draw inferences about their subjective well-being and preferences. [132]

The miscontrual problem refers to another mental process that may bring about dynamic inconsistency. It occurs because individuals make poor judgments about the objective attributes of particular objects or situations that elicit affective reactions (Loewenstein and Angner 2003, pp.367-368).

Individuals' difficulty in predicting how their preferences and feelings adjust to favorable or unfavorable situations give rise to inferential activities that sometimes produce

132 To investigate this source of preference misprediction, Schkade and Kahneman (1998) asked college students in California to estimate the happiness of Mid-Western students. They found that individuals in the Midwest believed that people from California were happier, and people from California expected Mid-Western students to be less happy. But their estimates deviated from actual reports: both groups expressed similar levels of happiness. We can explain this discrepancy between individual's predictions and actual reports by the fact that individuals have difficulty in estimating others' feelings and preferences, especially when they rely on a single element (e.g. weather conditions) to draw inferences about subjective well-being.

intertemporally inconsistent behaviors (ibid, p.368-370).[133] A fourth empirically grounded process by which intertemporally inconsistent behavior occurs has to do with the agent's limited ability to adjust her or his preferences to what she or owns (ibid). [134, 135]

Finally, the hot and cold empathy gap suggests that the agent experiencing a hot visceral state (e.g. hunger, sexual arousal, or fear) has difficulties in predicting what she (or others) will think, prefer, feel and do under the experience of cold visceral states (ibid, pp. 371-372) The inclusion of explanatory items that replace some elements of the previously isolated explanantia set - independence, time-consistent preferences, shifting preferences over time in predictable and non-predictable ways - promises to yield an improved representation of actual intertemporal decision-making that even describe those processes causally relevant to the occurrence of explanandum phenomena under study.

The empirical model of projection bias offers a straightforward account of intertemporal choice that accommodates some puzzling instances of behavior that are also covered by the anticipatory model and the visceral model (e.g. savoring effect, dread effect, dynamically inconsistent behaviors). Its comparative advantage lies in its capacity to

133 In order to test this empirical hypothesis, Pinel, Wilson, Blumberg and Wheatley (1997) asked subjects (assistant professors at University of Texas) to predict their future well-being at various points in time subsequent to a positive or negative decision over tenure. Their findings reinforce the hypothesis that individuals underestimate their own ability to adjust to (un) favorable situations and this process may be a source of mispredicted preference. Subjects reported that their preferences would be adapted to the tenure decision within a five-year period. However, assistant professors expected that their preferences would take from 5 or 10 years to adjust later stated that, regardless of a positive or negative decision, they had adjusted much faster than expected. In other words, they exaggerated the hedonic consequences associated with a tenure decision (promotion). In another experiment, Sieff and colleagues (Sieff et al 1998) found behavior patterns that confirm the preference adaptation hypothesis. They asked patients who were about to take a HIV test to estimate their feelings after receiving a negative or positive result. The experimenters stress that participants underestimated their capacity to adjust their preferences and feelings to a good or bad future outcome. Subjects exaggerated how happy they would feel on learning that they were HIV negative or how sad they would be on being told they were HIV positive.

134 Kahneman and Snell (1990) and Loewenstein and Adler (1995) found that people had difficulties with anticipating how their preferences would adjust to ownership. O'Donoghue, Loewenstein and Rabin (1999) suggest that an individual's difficulty in predicting her or his future preferences amplifies the endowment effect, and go on to explain how this apparent puzzle is produced. The typical buyer tends to overestimate the pleasures from acquiring a (rewarding) good, whereas the seller overestimates the negative utility consequences (pain) from losing a good. For details of the endowment effect, see Thaler (1991, 1992).

135 The model of projection bias offers an account of the endowment effect. This empirically observed phenomenon reveals that people value objects they own more highly than things they do not. Then, individuals are more sensitive to losing a good in their possession than gaining the same good. For more details, see Thaler (1991 and 1992).

describe mental processes needed for understanding of how certain patterns of behavior are produced. For instance, the model yields an explanation of intertemporally inconsistent behavior in terms of people's difficulty in predicting future utilities and outcomes (rather than by the desire for immediate gratification or higher levels of impatience).[136] It explains the puzzle of endowment effect in terms of an error caused by individual's reliance on current preferences in predicting future utilities and feelings (Loewenstein et al 2000, p.16).[137] In short, the model of projection bias performs better than other accounts that employ the method of de-isolation because it succeeds in isolating some emotional and cognitive processes with crucial roles in production of choice over time. In addition, the model has the capacity to unveil processes that do the explaining of why and how some people behave suboptimally under particular situations.

5. An explanation of intertemporal choice at two levels of analysis

Following up on chapter 4, this section discusses the applicability of a two-level account of decision-making behavior to the domain of intertemporal choice. The idea is to show that a model that offers a description of the decision machinery at the levels of physical brain and mind is worthwhile because it sheds extra light on the (causal) significance and robustness of emotional and higher-order cognitive processes underlying manifest intertemporal choice behavior. The proposed model is thought to add to behavioral economic accounts in terms of affect-based inferential procedures like "project one's current preferences, desires and feelings on future preferences, feelings and concerns". Unlike the model of projection bias, the two-level account does not view cue-based projection heuristics as necessary sources of biased or distorted behavior.[138] Rather, the

136 Loewenstein et al (2000) suggests: "in our model…time inconsistency in perceived preferences derives solely from misprediction of future utilities…the person is unaware of this inconsistency" (p.18)

137 The authors argue that projection bias explains amplification of endowment effect, i.e., individual's exaggeration of duration of perception of loss and sensation of gain. In their words, "because losses loom larger than gains, the exaggeration of losses has a greater impact (ibid, p. 16)

138 Loewenstein and his colleagues take projection as a source of cognitive error or visceral error. As they put it,

> Interpreted literally, our model suggests that a projection bias is a cognitive error. We assume the person attempts to maximize her intertemporal utility, but fails in this optimization due to a misprediction. An alternative interpretation is that the projection bias is a visceral error… wherein the person makes accurate predictions – or would readily do so – but is simply not attending to all the relevant utility changes when making decisions (Loewenstein, O'Donoghue and Rabin 1999, p.40)

model assumes that projection is a simple heuristic that – depending on the occasion – may lead to quick and adaptive inferences and choice or biased outcomes.

The two-level account hypothesizes that projection plays crucial roles in decision-making over time. This is because it involves mental processes (physically realized by particular brain structures and information processing activities) by which the agent can identify important intertemporal prospects, draw quick and (often) effective inferences about alternative courses of action and opt for an alternative that satisfies her or his perceived intertemporal preferences, desires and concerns. One distinctive trait of the proposed model of choice is that it offers an explanation of intertemporal decision-making that locates those emotional and higher order cognitive processes that are supposed to play major roles in the production of actual behavior.

5.1 Cerebral underpinnings of choice over time

Neuroscientists have suggested that particular brain structures have the capacity to mobilize and coordinate domain-specific information-processing activities that constitute the physical substrate for the human ability to make intertemporal judgment and decision-making (Camerer, Loewenstein and Prelec 2003; Manuck, Flory, Muldoon and Ferrel 2003). Their empirical findings have caught behavioral economists' attention since they inform economists about some isolated processes with an important role in the chain of events that bring about behavior patterns (DU anomalies included).

There seem to be four main brain structures that play major roles in the mobilization and coordination of information processing activities causally sufficient to the occurrence of choice over time: the prefrontal cortex; the amygdala; the hippocampus, and the nucleus accumbens.[139] The prefrontal cortex physically executes cognitively demanding mental processes causally relevant to the occurrence of choice over time, such as careful estimation of intertemporal prospects and the associated hedonic consequences. This is partly due to its connections with various regions of the brain (and the body overall). Based on studies of individuals with impaired prefrontal cortex, neuroscientists found

139 See chapter 4 for details.

that the prefrontal cortex physically executes significant and robust processes underlying intertemporal choice. Prefrontal cortex patients have difficulties in regulating those inferential processes that would enable them to make quick and effective decisions. Some neuroscientists suggest that the prefrontal cortex regulates information-processing responsible for production of real decisions in real time (Gazzaniga, Ivry and Mangun 1998; Adolphs and Damasio 2001). In addition, there is evidence that prefrontal cortex patients reveal systematically myopic preferences (i.e. they opt for immediate gratification) and cannot learn the (negative) utility consequences associated with previous myopic behaviors (Damasio 2003). The amygdala is another brain structure related to the individual's ability to make intertemporal judgments and decisions. As explained in chapter 4, the amygdala shapes choices by means of different channels. Given that this brain structure is the center of negative emotion-processing (e.g. fear), it coordinates processing activities that enable agents to exercise their mental capacity for quick detection of an important danger or opportunity posed to the environment and for preparing the agent to draw on previous experiences and their affective consequences in order to behave adaptively to it. [140] Together the amygdala and the prefrontal cortex coordinate processing activities that allow agents to allocate attentional, memory and learning processes so as to make judgments about available intertemporal prospects and the consequences associated with them (Damasio 2003). [141] There is evidence that amygdala-impaired patients have problems with coordinating those mental processes causally relevant to the occurrence of choice behavior. Emotionally flat agents (i.e. individuals with a damaged amygdala) have limited ability to learn from previous mistakes that were associated with negative hedonic outcomes and to draw on their feelings as sources of inferences about available intertemporal prospects to be pursued (or avoided). The hippocampus also serves an important role in the coordination of mental processes relevant to the occurrence of choice over time. It is responsible for the individual's mental capability for encoding, storing and retrieving information from previous experiences, and for long-term memory (LeDoux 1996/1998). There is evidence that individuals with impaired hippocampus have a reduced capacity to make those

140 LeDoux (1996/1998)

141 Ibid.

inferential activities, on which intertemporal choice depends. The nucleus accumbens is subcortical brain structure that is involved in positive emotion processing, reward and motivation. It mobilizes processing activities that enable an individual to ascribe positive affective values to particular intertemporal prospects and to identify one course of action within the perceived set of salient options that is thought to satisfy first her or his intertemporal aspiration level. The nucleus accumbens influences perception of individual preferences for certain intertemporal prospects and therefore may guide intertemporal decisions in a non-negligible manner (Mogenson, Jones and Yim 1980, Loewenstein, Read and Baumeister 2003).[142]

5.2 Detection, search and selection in the machinery for intertemporal choice

The suggested two-level account of intertemporal choice offers a description of how various components of a decision making system operate to produce manifest behavior. It shows that affect-driven processes play a central role. From this perspective, affective states and emotional processes are involved in the three information processing stages of a (theoretically described) decision-making mechanism, viz. detection, selective search and affect-driven selection.

5.2.1 The detection stage

The suggested two-level account of intertemporal choice differs from other behavioral models to the extent that it specifies mental processes (physically realized by particular brain structures with the capacity to undertake domain specific information processing activities) by which the agent can identify an important intertemporal choice task and mobilize her or his limited computational capabilities to respond adaptively. At this first information processing stage of the decision machinery, neural structures like the sensory thalamus, the amygdala and the sensory cortex coordinate activities through which the individual can identify and assess quickly the significance of the particular choice task at hand. Emotions play an important role in quick detection and appraisal of a decision problem situation to the extent that they yield credible cues about the significance of a

142 See Manuck et al (2003) and Cardinal et al (2001), for details of studies with animals in which damage to the nucleus accumbens is induced to study its impact on preference for immediate gratification.

choice task for the individual's survival and reproductive success (Cosmides and Tooby 1990). Emotions put a frame on agent's mind and therefore guide her or his judgments about the decision problem and its significance. Note that the argument developed so far indicates that emotions play an important role in quick detection of a choice task by virtue of the functions they serve. Yet the suggested account does not imply any commitment to the view that emotions always lead to quick and adaptive behavioral reactions (see chapter 4).[143]

5.2.2 Search stage

Inspired by Simon's theoretical model of boundedly rational choice, the two-level account of intertemporal choice suggests that this second stage of processing is physically executed by four mainly brain structures, viz. the amygdala, the prefrontal cortex, the basal ganglia and the hippocampus coordinate information processing activities that are responsible for the agent's capacity to draw effective inferences about the available intertemporal choice options and the consequences associated with them. This is partly so because these neural entities constitute trajectories through which the agent can draw on her or his own (and others') previous experiences (stored in memory and learning systems) that enable him or her to find alternative ways of dealing with a particular choice task. Provided that emotions serve the evolutionary cognitive function of quick mobilization and coordination of important atttentional, memory and learning processes by which individuals can find and evaluate intertemporal choice prospects (and their payoffs associated with them), they cannot be viewed as an elements playing minor (causal) roles in agent's selective search for satisfactory decision options.

5.2.3 Affect-driven selection stage

The two-level model of choice offers a description of the selection stage of the decision machinery that is compatible with Loewenstein's account of intertemporal choice in terms of projection heuristics. Yet the suggested explanatory account does not commit itself to the vision that projection is always a source of cognitive error. Rather, the model

143 According to Damasio (2003, p.147), "emotions and feelings have no crystal ball to see the future. Deployed in the right context, however, they become the harbingers of what may be good or bad in the near or distant future."

suggests that specialized brain structures like the prefrontal cortex, the amygdala and the nucleus accumbens coordinate processing activities by which the individual can exercise their mental capacity to draw inferences about the utility outcomes associated with emotionally salient options and to select an option within the choice set that she feels that first satisfies her perceived intertemporal aspirations and concerns.[144]

The above description of the decision machinery allows for an explanation of intertemporal decision making in terms of a sequence of specialized emotional and higher order cognitive processes (instantiated by specialized physical brain structures) playing major roles in the causal production of actual behavior. The briefly sketched two-level account yields a more general treatment of intertemporal choice behavior than the model of projection bias since it accommodates situations in which the agent's reliance on affect-driven projection heuristics to estimate available prospects (and the consequences associated with them sometimes) give rise to quick and adaptive outcomes and other times to dynamically inconsistent behaviors (Cosmides and Tooby 2000), such as preference reversals. The model implies the vision that, under conditions of extreme uncertainty about what the future may bring, projecting one's own and other's current experiences to predict future preferences and feelings can bring about effective intertemporal behaviors.

As I tried to show, the two-level account of decision-making contributes to the development of a genuine (causal) explanation of intertemporal choice behavior since it describes a mechanism supporting the mental processes (describable at the physical level as information processing activities undertaken by particular brain structures) that bring about intertemporal judgments and decisions. In such a perspective, it is worthwhile to describe the operation of the decision machinery at the physical brain level since it informs economists about the causal importance and the robustness of those emotional and higher order cognitive processes for actual choice over time. By specifying how

144 Following on the previous chapter, we can suggest that the agent's feelings offer cues to future hedonic consequences of prospects (i.e. courses of action). For instance, positive feelings inform her of future rewarding outcomes to be pursued, whereas negative feelings offer cues as to intertemporal options to be avoided (because they involve punitive consequences)

specialized brain structures mobilize processing activities by which individuals can exercise their abilities to detect a problem situation, to search for alternative intertemporal prospects and to select a course of action, the suggested model shows that described emotional processes are robust in the sense of guiding decision outcomes in systematic and non-random manner. In addition, a study of the decision machinery at the hardware level offers understanding of why emotions are significant for the occurrence of actual behavior. The two-level model shows that the quality of intertemporal choices is dependent on whether individuals are equipped with certain brain structures responsible for emotion processing among other activities (Damasio 1994, 1999 and 2003, Adolphs and Damasio 2001).

The suggested two-level model is in tune with Loewenstein et al's vision that various patterns of intertemporal choice behavior can be better by making reference to an affect-driven projection heuristic (bias). It also puts forth the idea that emotions, feelings among other affective processes and states play major roles in production of some behavior patterns regarded as anomalies. But unlike other behavioral models of intertemporal choice that take emotions seriously, the two-level account offers a way of specifying how emotion, higher order cognition and manifest choice behavior are related in a complex economic world of causal (inter)dependencies.

The two-level treatment of intertemporal choice has the advantage of uncovering information-processing activities (performed by specialized brain structures) that constitute the physical foundations of the human capacity to coordinate affective and cognitive processes with major roles in actual behavior over time. The proposed model sheds further light on the significance of some mental processes (e.g. hot and cold empathy-gap among other sources of affect driven inferential activities) for the occurrence of actual behavior patterns (some DU anomalies included). An explanation of intertemporal choice at the levels of mind and brain shows that projection is a significant heuristic in a world where decisions about future prospects are needed. In this context, a two-level account attempts to contribute to a clearer picture of choice over time and its apparent puzzles.

6. Concluding Remarks

Choice over time is a pervasive kind of economic behavior in need of a clearer understanding. This chapter contributes to accomplishing this task by looking at various models of choice developed to yield improved predictions and explanations of behavior patterns relative to the mainstream approach to intertemporal choice (e.g. discounted utility framework).

Driven by the idea that economists are worried about the predictive and explanatory implications of their models, I discussed some empirical anomalies that prompt economists to question whether the standard account of intertemporal choice had excessively or incorrectly isolated those explanantia factors playing key roles in the phenomenon under study. I showed that doubts about the sufficiency of the isolated explaining items to an adequate account of actual choice over time (that meets the scientific purposes of accurate prediction and explanation in a richer causal sense) gave rise to alternative models that pursue two reformist strategies – incremental analysis and process-description strategy. The incremental analysis can be regarded as a de-isolative move since inclusion of new explanatory items occurs so as to supplement the previously isolated explanantia set (that failed to deal with important explananda phenomena), while the process-description strategy seems to be a re-isolative move because inclusion of new explanatory items happens so as to replace explanatory items that from the previously isolated set of explainers. Following up on chapter 3, I discussed behavioral models of intertemporal choice that relax some behavioral assumptions that served to exclude explanatory items that are now thought to play significant (or at least non-negligible) causal roles in production of behavior patterns. I also showed what implications reformed models of intertemporal choice carry to progress as scope expansion and enhanced causal articulation.

In this chapter, I paid special attention to a model of projection bias that uncovers processes by which intertemporal anomalies happen in the world. One novel feature of the projection model is that it converges with a heuristic approach to decision-making behavior and gives an improved understanding of the role of affective processes in the

production of actual choice over time. The model of projection bias involves the notion that individuals show dynamically inconsistent preferences because they estimate future preferences, goals and concerns based on the projection of their current affective states. Its gain in explanatory power lies in the model's ability to isolate affective and cognitive processes with important roles in the (causal) production of choice over time (from those that are negligible or unimportant).The model is good at predicting intertemporal choices that remain a puzzle to the DU framework, such as consumption and saving behaviors that reveal the phenomenon of 'hot and cold empathy gap'. By so doing, the model adds to explanatory progress in terms of causal articulation. Nevertheless, the model involves the view that one's projection of current preferences (i.e. feelings, goals, values) onto the future always explains time inconsistent behaviors. This is a controversial claim to the extent that the model cannot accommodate situations in which reliance on an affective-driven mental shortcut (such as a projection heuristic) gives rise to quick and effective intertemporal judgments and decisions. Furthermore, the model of projection bias does not sufficiently specify how affective and cognitive processes interact to bring about actual intertemporal choices. To circumvent this limitation, I proposed a multi-level account of intertemporal choice that purports to be a development of Loewenstein and his collaborators' ideas on the driving forces behind actual decision-making over time. The proposed account of intertemporal choice includes a brain level of analysis that throws extra light on the significance of cognitive and affective processes underlying choice behavior.

The first conclusion of this chapter is that most research in behavioral economics can be interpreted as de-isolative moves that relax some idealizing assumptions like constant discounting and positive time preferences so as to supplement the previously isolated set of explainers with some previously excluded explanatory items like intertemporally conflicting preferences and farsighted preferences with major effects on manifest choice behavior without removing. The incremental analysis can be regarded as a de-isolative move since inclusion of new explanatory items occurs so as to supplement the previously isolated explanantia set (that failed to deal with important explananda phenomena). However, some models like the reference-point model and the model of projection bias

pave the way for re-isolative moves, since they relax some assumptions like time consistency, independence, stationarity of instantaneous utility function so as to replace some previously included explanatory items of the isolated explanantia set (time-consistent preferences, preferences for certain intertemporal prospects regardless of the agent's past and future consumption profiles) with excluded explanatory items (e.g. time conflicting preferences, the agent's tradeoffs between consumption of certain goods is sensitive to her or his past and future consumption of these goods). In addition, these re-isolative moves seem to be substitute the account of choice over time in terms of its underlying (causal) processes for an explanation that refers to the hypothesis of discounted utility maximization.

The second conclusion is that a two-level description of the machinery for intertemporal choice contributes to the task of tracking the causal chain of states, events and processes that produce choice over time, partly because it points to the various ways in which emotion and higher order cognition interact to guide behavior. The two-level description of the decision machinery offers understanding of why judgments of intertemporal prospects and their hedonic consequences necessarily involve some affect-driven projection heuristics.

A third conclusion of the chapter is that an account at mental and neural levels does not imply any commitment to the idea that projecting one's current states onto the future necessarily results in intertemporal inconsistency. Rather the model suggests that one's reliance on projection heuristics sometimes lead to fast and effective judgments and decisions, while producing distorted outcomes on other occasions. The model suggests that there will be a biased outcome if judgments about intertemporal prospects and their hedonic consequences are associated with strong (instead of moderate) emotional reactions and feeling states. Of course there is a burden of the empirical proof involved here. There is evidence that agents under the impact of intense affective states tend to misinterpret the task at hand and to mispredict the nature, intensity, duration and effect of their current affective states on their own and others' future goals, preferences and feelings. This evidence should lead us to be conditional rather than deny that there are

situations in which reliance on gut feelings leads to quick and adaptive choices in a decision environment permeated by extreme uncertainty as to what the future may hold. For instance, contemporary neuroscientists emphasize that amygdala-impaired patients are emotionally 'flat' and have difficulty in making effective intertemporal choices in so far as they are limited in their ability to regard affective processes and states as significant cue-based heuristics through which agents quickly estimate the utility consequences of prospects and set criteria for satisfactory responses.

The fourth conclusion is that our description of the intertemporal decision machinery points to processes that enhance our grasp of how economically relevant DU anomalies occur, such as savings after retirement, preference reversals and so forth. In addition, a two-level approach fills a gap in the explanation of intertemporal choice by identifying the significant roles of emotions and feelings in judgment and decision-making. For that reason, I am led to suggest that a two-level approach might contribute to the behavioral economists' pursuit of explanatory progress as increased causal articulation and scope expansion.

CHAPTER 6
EMOTIONS IN THE EXPLANATION OF DECISION UNDER RISK

> Most, probably, of our decisions to do something positive, the full consequences of which will be drawn many days to come, can only be taken as the result of animal spirits – a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities… human decisions affecting the future, whether personal or political or economic, cannot depend on strict mathematical expectations, since the basis for making such calculations does not exist … it is our innate urge to action that makes the wheel go around.
>
> John Maynard Keynes, *The General Theory of Employment, Interest and Money*

> On the contrary, emotion probably assists reasoning, especially when it comes to personal and social matters involving risk and conflict. I suggested that certain levels of emotion processing probably points out to the sector of the decision-making space where our reason can operate most efficiently
>
> Antonio Damasio, *The Feeling of What Happens*

1. Introduction

Most (if not all) instances of human judgment and decision-making are made under conditions of risk or uncertainty. [145, 146] Individual and social behaviors with economic relevance often involve uncertainties as to the available prospects and associated (hedonic) consequences. Thus it is hardly surprising that the task of explaining risky

145 The great economist Frank Knight (1921) saw risk as referring to situations in which individuals know the distribution of possible outcomes, whereas uncertainty refers to cases in which prospects cannot easily be represented by means of a (objective) probability distribution function. According to him,

> The practical difference between the two categories, risk and uncertainty is that in the former the distribution of the outcome in a group of instances is known (either through the calculation a priori or from statistics of past experience), while in the case of uncertainty this is not true, the reason being that it is impossible to form a group of instances, because the situation dealt with is in a high degree unique (*Risky, Uncertainty and Profit*, Part III, Chapter VIII, paragraph 2)

146 The conceptual difference between risk and uncertainty loses part of its bite when choice theory draws on a subjective probability approach.

choice behavior (and its empirical puzzles or anomalies) has inspired much of research within economics and psychology. This chapter puts forth the idea that behavioral models of risky choice have been developed in response to decision theorist's doubts about whether the standard expected utility framework (henceforth: EU) selected a set of explanatory factors (by means of its underlying assumptions) that are sufficient or even necessary for a satisfactory explanation of patterns of risky choice behavior in the real world.

Following on from chapters 4 and 5, the current work examines some models of choice that have made some de-isolative or even re-isolative moves so as to develop improved theoretical representations of actual decision-making under conditions of risk or uncertainty that in turn meet the economist's purposes of explanation (prediction) of actual behavior. One of the main contributions of the chapter is to show that these reformed models of choice increasingly recognize the major roles emotions and feelings play in thinking and decision-making under conditions of uncertainty about available prospects and the consequences associated with them; and therefore they might contribute to explanatory progress as scope expansion as well as enhanced causal articulation.

Furthermore, the current chapter advances the thesis that a two-level account of decision-making under risk is worth developing to the extent that it offers a genuine (causal) explanation of risky choice that uncover those emotional and higher order cognitive processes playing major roles in production of behavior patterns, including some regarded as expected utility anomalies.

The overall argument is assembled in the next five sections. Section 2 briefly presents the main assumptions underlying the standard expected utility (EU) framework and the explanatory items that they seem to remove from the set of explainers (i.e. explanans set). It discusses two amended EU models viz. regret theory and disappointment theory that were developed to supplement the previously selected set of explanatory items from the EU framework with removed explanatory items that play major roles in actual risky

choice behavior. I go on to argue that both models – regret theory and disappointment theory- add affective factors (e.g. regret, rejoice, disappointment and elation) to the EU framework's *explanantia* set with the aim to resolve economically significant risky anomalies, such as certainty effect, preference reversal, simultaneous gambling and insurance and so forth. Section 3 examines an empirical hypothesis called 'risk as feelings' that makes early re-isolative moves in direction of an explanatory account of decision under risk that unveils affective and cognitive processes significant for actual behavior (Loewenstein, Weber, Hsee and Welch, 1998/2001). It shows that this empirically grounded model reveals the links between (a) risk perception and affective factors; (b) feeling, thinking, and probability judgment, and (c) experienced emotions and (utility) value ascription. Section 4 analyzes an account of risky choice called affect heuristic model that offers a second re-isolative move towards specification of major driving forces behind actual inferences and choices under risk (Finucane, Alkahami, Slovic and Johnson, 2000 and Slovic, 2002b). I claim that the affect heuristic model deviates more clearly from the EU framework and contribute further to explanatory progress in economics since it dispenses with the hypothesis of expected utility maximization and engages in describing those emotional and higher order cognitive factors that give rise to actual choices under risk/uncertainty.[147] Section 5 proposes a theoretical model of decision-making behavior with two levels of analysis that adds to the development of an affect-driven heuristic explanation of risky choice behavior. I go on to suggest that a two level account of risky choice behavior is worth developing to the extent that it uncovers causally significant emotional and higher order cognitive processes for the occurrence of manifest behavior. Section 6 amalgamates the chapter's main points and offers some conclusions.

2. Amended expected utility models that take emotions seriously

In chapter 2, I suggested that Daniel Bernoulli (1738)'s efforts to resolve an apparently empirical anomaly called the St Petersburg paradox led him to offer the first expected

147 Opening black boxes does not necessarily imply improvements in explanatory power. The gain in explanatory power depends on the model's capacity to identify processes or mechanisms (causally) significant for actual behavior. The explanatory improvement that I have in mind refers to increased causal articulation and scope expansion (though I must admit these are two different readings of progress). For details, see chapter 4.

utility theory of choice. Mathematical economists of the 20th century like von Neumann, Morgenstern (1944/1947) and Savage (1954) later worked on the development of an expected utility theory that ensured the correspondence between the required cardinal utility scale and an ordinal utility function (i.e. it reveals individual's preferences over gambles). Due to some theoretical difficulties associated with the vNM utility approach, Savage's subjective expected utility theory became the mainstream account of decision under risk/uncertainty.[148] Savage's framework consists of a set of states of the world (denoted by S), a set of consequences (denoted by C) and a set of acts or courses of action denoted by F. In his perspective, the agent's evaluation of a course of action in the face of risk (uncertainty) depends on the agent's taste for the possible consequences and her or his beliefs concerning their occurrence. His subjective expected utility theory postulates a preference structure that allows for a numerical representation of the individual's valuation of consequences by a utility function. The agent's beliefs about the occurrence of certain events are represented by a probability measure on the set of events. Then, the agent's evaluation of prospects can be given by mathematical expectations of the utility relative to the subjective probabilities associated with them.
Savage's account of decision under risk is derived from the following set of assumptions, denoted by SV:

SV1. Completeness and Transitivity. The agent's preference ordering is complete. For any two acts **a** and **b**, "**a** is weakly preferred to **b**" or "**b** is weakly preferred to prospect **a**". For all acts **a**, **b** and **c**, if "**a** is weakly preferred to **b**", and "**b** is weakly preferred to **c**", then "**a** is weakly preferred to **c**".
SV2. Continuity. For all prospects **a**, **b** and **c**, where "**a** is weakly preferred to **b"** and "**b** is weakly preferred to **c**", there exist some probability **p** such that a "compound prospect (**a**, **p; c, (1-p**)) is indifferent to an act **b".** As a result, there exists a function **V (.)** which ascribes a real-valued index to each prospect such that **V(a) ≥ V (b)**, if and only if, "an act **a** is weakly preferred to **b"**.
SV3. For every event E and acts **a** and **b**, "**a** is weakly preferred to **b"**, if and only if, "a compound prospect **(a, p, c (1-p))** is weakly preferred to a prospect **(b, p; c (1-p))**" for every act **c**

148 For details, see chapter 2.

and there is a utility function showing the consequences associated with available prospects (independence of consequences)

SV4. Sure Thing Principle: for all acts, a, a', b, b' and every event E, the prospect "(a, p; b(1-p)) on event E is weakly preferred to **(a', p; b(1-p))**" if and only if "prospect **(a, p; b'(1-p))** on event E is weakly preferred to **(a', p; b'(1-p))**". In this case, preference over acts only depends on the consequences in states in which the payoffs of the two compared acts are distinct.

SV5. The agent's preferences over certain prospects reveal the subjective probabilities ascribed to them (i.e. beliefs about their likelihood)

SV6 For all events E and E' and prospects **a, a', b, b'** such that "**a** is strongly preferred to **b**" and "**a'** is strongly preferred to **b'** ", we can say that "compound prospect **(a, p, b (1-p))** on event E will be weakly preferred to a prospect **(a, p; b (1-p))**" for event E', if and only if, "compound prospect **(a', p; b',(1-p))** for event E is weakly preferred to a compound gamble for the event E' **(a', p; b'(1-p))**"

SV7 For some prospects that give the same consequences (i.e. constant acts), it can be said that a certain "act **a** is strongly preferred to another **b**"

SV8. For all acts **a**, **b** and **c** satisfying the relation "**a** is strictly preferred to **b**", there is a finite partition of events [denoted by $(Ei)^{n}_{i=1}$] from the set of states such that, for all i, " **a** is strongly preferred to **(a, p; c, (1-p))**" on the event E_i and "**(c, p; a (1-p))** on event E_i is strongly preferred to an act **a**".

SV9. Monotonicity. For every event E and all acts **a** and **a'**, if "**a** is strictly preferred (on event E) to **a'**(s)" for all states s in E, then "**a** is strongly preferred (on the set of events, E) to prospect **a'** "

Based on this set of assumptions, Savage found out that there exist a unique and finitely additive probability measure, denoted by **p** on the set of states, S and a real-valued function **u** on the set of consequences unique to a positive linear (affine) transformation such that an act **a** dominates another prospect **b** if $\sum_{j=1}^{n} p.a_j \geq \sum_{j=1}^{n} p\ b_j$. These ideas gave rise to the subjective expected utility theory, which became the 'crowning glory' of rational choice theory. Savage's assumptions serve to remove some explanatory items that were regarded to play negligible roles (minor causal roles) in actual decision under risk. To better understand the chosen set of explainers that characterize the subjective utility approach (later criticized by behavioral researchers), it may useful to highlight some of the functions of certain SV's. For instance, SV1 and SV2 seem to be applicability assumptions; these idealizations serve to remove schemes of preference ordering incompatible with formal

derivation of a well-behaved utility function that represents agent's valuation of prospects. SV3 (independence of consequences) serve the role of negligibility assumption, since it removes an explanatory element (i.e. individual ordering of consequences is dependent on the event and the act associated with it) that is thought to play negligibly small effects on actual decision under risk. This assumption also helps one to derive an analytical model in which an individual's belief that an event E is at least as likely to obtain as another event E' reveals her or his preferences for a bet on event E to the same bet on E' (this is one of the assumptions that many behavioral models of risky choice relax). Related to that, the sure thing principle (SV4) is a negligibility assumption that removes 'preferences that are dependent on other things than the consequences in states of nature in which the payoffs of the two compared prospects (gambles) are identical', since they are thought to play non-important roles in patterns of actual behavior. SV5 seems to be a negligibility assumption that removes the factor that people's preferences over certain prospects might not reflect their beliefs about the likelihood and outcomes associated with them. SV6 seems to be a negligibility assumption that removes the existence of event-dependent attitudes towards risk. SV7 is a negligibility assumption that serves to exclude the factor 'individual can be indifferent among all acts'. SV8 seems to be a domain assumption that is used for analytical convenience, i.e. it excludes a non-negligible factor 'individuals can regard a consequence infinitely better or worse than any other consequence' that needs to be absent for derivation of a well-behaved probability measure representing the agent's beliefs in the various (infinite) states of the world. SV_9 (monotonicity) is an idealizing assumption that seems to make a model of subjective expected utility applicable. It implies that 'for a decision maker that considered an act strictly better (or worse) than payoffs associated with another act, the former prospect is conditionally strictly preferred (or strongly less preferred) than the latter". As a result, a (subjective expected) utility model can be applied to explain (predict) a wide range of choice phenomenon (see Friedman and Savage 1952; Fishburn 1970, Kreps 1988). The following table summarizes the main SEU assumptions and what explanatory factors that it removes.

Main assumptions of the SEU model	What explanatory factors it removes
Complete and Transitive preference ordering	They exclude the existence of preferences over prospects that are not well-behaved
Continuity	Existence of preferences that cannot participle in an indifference zone and therefore cannot be represented by a real-valued function that serves as an index of risky prospect ranking
Sure thing principle (type of independence)	Preferences that are dependent on other things than the consequences in states of nature in which the payoffs of the two compared prospects (gambles) are identical'
Individual ranking of consequences is independent of the event and act that give rise to it	People order consequences depending on the events and prospect that yield them. People's beliefs about the occurrence of an event do not reveal their preferences for a prospect on this state
People's beliefs about the likelihood of occurrence of behavior reveals their preferences for the prospects on certain states	Individual beliefs about occurrence of a certain consequence is independent from her preferences among acts
Betting preferences are independent of the specific consequences that characterize the bets	Individual's attitudes towards risk are event dependent
No consequence can reverse a strict preference ordering of two acts	Consequences can be perceived as infinitely better or worse than any other consequence
Monotonicity	Agents' perception of an act that is strictly better (or worse) than payoffs associated with another act do not mean that the former act strictly more preferable (or strongly less preferable) than the latter

TABLE 6.1

List of main assumptions underlying the subjective expected utility model and the explanatory items that are removed from the explanantia set

Some expected utility anomalies like Allais' and Ellsberg's paradoxes (empirical violations of SV_4 and SV_5) led decision theorists to doubt whether the isolated

explanatory factors by the main assumptions underlying the subjective expected utility framework ought to be supplemented or even replaced by previously excluded elements.

This particular section takes up regret theory and disappointment theory as the first amended expected utility models that explicitly take emotions as important factors to the occurrence of actual decisions under risk. The next lines show that both analytical models are built on the presumption that anticipated emotions (e.g. regret, rejoice, disappointment and elation) guide human inferences as well as actual choices under conditions of risk/ uncertainty in relation to prospects.

2.1 Regret theory

The economists Graham Loomes and Robert Sudgen (1982) provided regret theory (see also Bell 1982). The latter is here interpreted as a reformed expected utility model that arises in response to doubts about whether some previously excluded explanatory factors from the set of explanantia of the SEU model play major (rather than minor) roles in production of economically relevant risky choice behaviors. Regret theory seems to question the sufficiency of the isolated explanatory items by the SEU model for a satisfactory account of empirical regularities of choice behavior under conditions of risk.

Loomes and Sugden develop a model of risky choice that incorporates expected emotions (e.g. regret, rejoice) into the *explanantia* set of the EU framework with the aim to explain (predict) actual risky choice behavior including those behavior patterns viewed as anomalies. The reformed DU model hypothesizes that "one significant explanatory factor [affecting people's risky choices] is an individual's capacity to anticipate feelings of regret and rejoicing" (Loomes and Sugden 1982, p.822). To incorporate the above explanatory item, they relaxed assumptions like independence and sure thing principle.

It is important to stress how emotions are formally treated in regret theory. The de-isolated model assumes that individual choices are based on comparisons between the expected consequences of alternative options. Regret theory defines risky preferences and decisions in relation to a pair of potential acts. On this view, any two acts A_i and A_j are

assumed to yield outcomes (consequences), x_{is} and x_{js} in the state of nature S. The individual's utility of x_{is} is represented by a preference function M (x_{is}, x_{js}), where $\delta M/\delta x_{is} > 0$ and $\delta M/\delta x_{js} < 0$.

Given that Loomes and Sugden see utility as having an explicit affective foundation, they assumes that an agent's perception that she or he satisfied a particular desire will elicit particular emotional experiences (ibid, p.807). These ideas are formally represented as follows. In cases where $x_{is} < x_{js}$, M (.) is associated with feelings of regret. When $x_{is} > x_{js}$, the function M (.) yields feelings of rejoicing. This is partly so because an individual anticipates intense feelings of regret at the perception of a huge difference between what she or he actually gets from a chosen prospect and what could have been received had she selected any other (foregone) prospect.

Regret theory seems to follow an incremental strategy of theorizing to the extent that it assumes that a rational decision maker aims to select a prospect that yields minimum regret feelings and maximizes $\sum_s p_s M (x_{js}, x_{is})$, where p_s corresponds to the probability of occurrence of state s.[149] Then, it can be read as an de-isolative move that relaxes assumptions of independence so as to come up with an improved theoretical representation of actual decision under risk, without questioning the core hypothesis of expected utility maximization. Yet some ideas behind regret theory differ in interesting ways from standard EU framework. The model uncovers some psychological determinants of actual risky choices, anomalies included. The merit of regret theory lies on its capacity to deal with anomalous patterns of behavior, such as the reflection effect, overweighing of small probabilities, common consequence effect (certainty effect), or simultaneous gambling and insurance (Bell 1982, Starmer 2000).

Reflection is a type of framing effect, i.e., a behavior pattern that cannot be explained (predicted) by standard EU theory (Kahneman and Tversky 1979). It is observed when an individual shows preference reversals between gains and losses. For instance, agents that are risk averse in the domain of gains (that is, they prefer a lottery with a certain win of

149 Expected utility model is a special case of regret theory when M(x is , x js) = U(xis).

$3,000 to one where they have an 80% chance of winning $4,000) become risk seeking in the loss domain (a lottery with a 80% chance of losing $4,000 is preferable to one with a sure loss of $3,000). Regret theory gives a straightforward account of the reflection effect. It suggests that individuals, willing to minimize regret feelings, prefer a lottery with a certain win (of $3,000) to one with a 20% chance of winning nothing (and an 80% chance of winning $4,000). Quite analogously, individuals that aim to minimize regret will prefer a lottery associated with 80% probability of losing $4,000 and 20% chance of no loss to a sure loss of $3,000. This apparent puzzle is resolved once decision-makers are expected to feel awful at the perception that they opted for a sure loss instead of a lottery with no loss.[150]

Regret theory can explain why individuals sometimes select a lottery with high probability of gaining a small amount of money and at the same time ascribe a high price to a gamble involved with small probability of making a large profit. The above case of preference reversal is regarded as an EU anomaly. It shows the possibility of individual's overweighting of small probabilities. Regret theory accommodates the phenomenon by reference to the agent's anticipation of regret feelings (resulting from by her or his comparison between chosen and foregone prospects). From this perspective, expected emotions motivate the agent to prefer a lottery with high probability of winning a small amount against a gamble with a low probability of winning a large sum.[151]

Regret theory also provides an account of a famous EU anomaly called Allais' paradox (also known as common consequence or certainty effect), already cited in chapter 2.[152] Consider two problem situations. In the first you have to choose between gaining $1 million for sure, L_{11}= ($1mi, 1) and participating in a lottery where the chances of winning $5 million are 10%; $1 million 89% and nothing 1%, L_{12}= ($5mi, .1; $1mi; .89; $0; .01). The second problem asks individuals to select between one lottery that yields $1million with 11% of chance or nothing with 89% denoted by L_{21}= ($1mi, .11; $0,.89)

150 Kahnmenan and Tversky explained the reflection effect by suggesting that certainty enhances the aversion to losses and the desirability of gains (1979, p.265).
151 A lottery associated with small probability of winning a large amount is more conducive of intense regret feelings.
152 For details, see Allais (1953).

and another that yields $5million with 10% probability or nothing with 90%, L_{22}= ($5mi, .10; $0,.90). The expected utility approach assumes that individuals may select the pair (L_{11}, L_{21}) or (L_{21}, L_{22}). Empirical evidence challenges this prediction, since individuals that chooses L_{11} over L_{12} due to the certainty of gaining a large sum can also select L_{22} (instead of L_{21}). This is because the odds of winning are very similar. According to Bell (1982), regret theory can resolve the above preference reversal. In his words,

> A decision-maker who takes the gamble over a sure gain of $1 mi may feel absolutely devastated if the 1% chance of getting nothing results. It would be considered normal, if not economically rational, for such a decision-maker to feel angry and perhaps depressed that he or she could have been so stupid or greedy as to pass up a sure $1 million… In the second choice between gambles there is no equivalent endpoint. If the chosen lottery results in no reward, the decision maker may feel that nothing was also the likely result of the other lottery. (1982, p.962)

Regret theory can predict other puzzles to the body of EU framework such as simultaneous gambling and insurance (Friedman and Savage 1948) that reveal the constraint put by assumptions like independence and sure thing principle. Regret theorists suggest that individuals are partly ready to participate in a lottery in order to minimize the regret they would experience if they had chosen not to make a bet they would have won. They are also willing to buy an insurance against a hazard with small probability because they anticipate the regret feelings they would experience had they chosen not to be insured against a large loss that actually occurred. This amended version of the expected utility framework is a theoretical improvement that covers behavior patterns covered by the EU framework as well as certain behavior patterns regarded as risky choice anomalies. In this respect, it may pose an interesting alternative to mainstream choice theory since it contributes to predictive improvements (Sudgen and Loomes, 1986). Nevertheless, regret theory also faces limitations. It has constrained explanatory capabilities, particularly in that it is a model of pair-wise choice under uncertainty and therefore cannot have the same broad scope as expected utility theory (Quiggin, 1994; Starmer, 2000). On this reading, regret theory may not be interpreted as an actual alternative to the expected utility approach. One source of objection to regret theory is that it hypothesizes that two expected emotions - regret and rejoicing mainly shape judgment and decision under risk or uncertainty. A second criticism is that regret theory also isolates too much, since assumes that other ways in which regret and rejoicing shape

the agent's perception of the choice task and her or his perception of risk are unimportant (Keasey, 1984). A related source of objection is that regret theory downplays the significant role some immediate emotions (e.g. fear and anger) among other affects play in an explanation of decision making under risk and uncertainty (Loewenstein and Lerner, 2003).

2.2 Disappointment theory

In response to criticisms of regret theory, Bell (1985), Loomes and Sugden (1986) and Loomes (1987) came up with 'disappointment theory' - another effort to de-isolate the previously chosen explanantia set that characterizes the SEU model. While regret theory assumes that risky decisions are based on counterfactual comparisons of two outcomes within the same state of the world, disappointment theory is distinguished by the assumption that actual choices depend on comparison of outcomes in different states of the world. This amended version of the expected utility model assumes that feelings of disappointment or elation are triggered in the agent at the perception that his or her selected gamble yields less or greater hedonic outcomes than expected. [153] Like regret theory, disappointment theory assumes that a prospect can be represented by an action with consequences contingent on the possible states of nature. This amended model of risky decision assumes that an individual forms prior expectations of a prospect based on its expected utility value, $\underline{U}$. Note that disappointment theory also seems to result from a theorizing strategy of incremental analysis. It gives a formal treatment to emotions like disappointment or elation by specifying a function D (.) that represents the net utility resulting from the difference between the actual consequence of a prospect (after uncertainty is resolved) and its expected utility value. This analytical account of risky choice hypothesizes that an individual act as if she or he behaved to maximize the following modified expected utility function:

$$E_i = \sum^{n}_{s=1} p_{is} [U (x_{is}) + D (U (x_{is}) - \underline{U})]$$

153 According to Loomes and Sugden, "when considering any uncertain prospect, an individual forms some prior expectation about that prospect: if the consequence falls short of the prior expectation, then in addition derived from the consequence itself, the individual also experiences some degree of disappointment; whereas if the consequence is better than the prior expectation, the individual feels some measure of elation" (1986, p.271).

When D (.) = 0, disappointment theory is equal to the standard expected utility model. Similarly to regret theory, this variant of the expected utility model adopts a consequentialist perspective that views emotions and feelings as affective experiences that individuals anticipate in order to maximize their expected outcomes. Disappointment theory can predict behavior patterns covered by the EU approach and also behavior patterns that are viewed as anomalies, such as the Allais paradox, overweighting low probabilities and underweighting high probabilities, simultaneous gambling and various violations of the compound probability axiom. This de-isolative move allows for a model of risky choice that explains (predict) the above empirical violations by referring to the agent's limited capacity to anticipate disappointment (or elation) feelings on perceiving the outcome of an action that turn out to be less or (greater) than expected.

Disappointment theory and regret theory have similar structures. Both modify the basic EU model to enhance the quality of predictions of decision under risk. According to Loomes and Sugden, regret theory and disappointment theory assume that individuals seek to maximize a mathematical expectation of utility defined as hedonic satisfaction. Both depend on comparisons of outcomes associated with available actions.[154] Loomes and Sugden stress that regret theory and disappointment theory do not provide mutually exclusive (i.e. competing) accounts of decision behavior under risk or uncertainty. Disappointment theory can explain certain behavior patterns that regret theory accommodates (including anomalies such as the common consequence effect, common ratio, overweighting small probabilities, simultaneous gambling and insurance). This being the case regret theory and disappointment theory may complement one another in developing an alternative theory of rational choice with improved explanatory and predictive powers (p.271).

The following sections take up these challenges posed to amended EU models and discuss some theoretical representations of decisions under risk that deviate more

154 The utility or satisfaction derived in disappointment theory may depend on comparisons between the consequence of a selected course of action and the other possible outcomes of this same action in other states of the world. In disappointment theory, utility or expected satisfaction results from comparison between the consequence of the action and the consequence of other actions within the same state of nature (Loomes and Sugden 1986, p. 281).

radically from the expected utility framework than regret theory and disappointment theory. The idea is to seek behavioral models that describe processes or mechanisms by which actual risky judgment and decision-making take place. In so doing these alternative models allow for re-isolative moves in behavioral theorizing on decision under risk that increasingly take emotions as playing major (causal) roles in production of economically relevant patterns of behavior and purport to replace the explanatory principle of expected utility maximization with affect-driven heuristics for judgment and decision-making behavior in the real world.

3. The 'risk-as-feeling' hypothesis and the theorizing strategy of re-isolation

Contradicting a long-standing tradition that accounts for risky choices in terms of purely cognitive activities, Loewenstein, Weber, Hsee and Welch (2001) suggest that emotions and feelings play an important influence in the operation of the decision machinery. They hypothesize that emotions and feelings are not epiphenomenal but mental processes playing crucial roles in individual's risk perception, probability judgments and estimation of utility outcomes associated with risky prospects. With that in mind, they claim:

> The risk as feelings hypothesis suggests that feelings play a much more prominent role in risky decision-making than they are given credit for by the cognitive-consequentialist tradition (…) Thus feelings may be more than just an important input into decision-making under uncertainty; they may be necessary and, to a large degree mediate the connection between cognitive evaluations of risk and risk-related behavior (Loewenstein *et al*. 2001, p.274)

The above ideas of developing an explanation of choice behavior in terms of the 'risk-as-feeling' conjecture draw on Paul Slovic's psychometric studies[155] of the emotional and cognitive dimensions of risk (Slovic, 1987 and Slovic and Peters, 1996). From this standpoint, risk comprises more than hazards represented by an objective distribution of possible outcomes and in fact has an affective foundation with a cognitive counterpart. The former component of risk is called dread risk factor, whereas the latter refers to an unknown-risk factor. Uncontrollable, potentially harmful or catastrophic hazards are associated with a dread risk component. Hazards that involve unobservable or unknown

155 Psychometrics is an empirical methodology that uses psychophysical scaling and multivariate analysis techniques in order to produce quantitative representations of risk attitudes and risk perceptions. For details, see Fischhoff, Slovic, Lichtenstein, Read and Combs (1978) and Slovic, Fischhoff and Lichtenstein (1984).

distribution of probabilities and outcomes with deferred results will score high on the unknown-risk component. In this case an account of behavior based on the hypothesis of risk as feeling needs to relax the assumptions of independence, sure thing principle, the vision that risk attitudes are independent from events and individual's ranking of consequences that occurs independently from one's perception of events and possible acts (risky prospects).

3.1 The emotional basis of risk perception and risky preference

Although decision researchers often study risk in terms of the likelihood of a hazard and its associated outcomes, the concept of risk is itself used in an obscure fashion. Slovic and Weber (2002) saw risk as variously equated with: (i) a hazard, (ii) a probability, (iii) a consequence and (iv) a significant danger or adversity. Slovic and his colleagues acknowledge that even the risk-as-feelings hypothesis belongs to this tradition of conceptual ambiguity. The novelty of this alternative account is that it does not imply a commitment to the partial interpretation of risk as mere affective states or feelings; nor does it deny that risk also has important cognitive roots.[156] The risk-as-feelings account assumes that emotions and feelings guide risk perception (defined in terms of the severity of outcomes that some prospects yield), preferences in relation to risky prospects, value judgment and choice behavior (Loewenstein et al 2001, p.271).

3.1.1 Affect and risk perception

The psychologist Alice Isen and her collaborators were the first to offer evidence that affective states shape the decision machinery in non-random ways due to their capacity to guide agent's perception of risk. More specifically, they designed experiments to investigate the impact of positive affective states on individuals' risk perception and their preferences for risky prospects. Based on a gambling task in which some participants were induced to experience positive affect, they discovered that positive-affect agents systematically preferred a small sure gain ($1) to a gamble with a high payoff ($10) with only a 50% chance of winning. These findings led to the conclusion that individuals

156 Risk also refers to the distribution of possible outcomes associated with a perceived hazard.

experiencing a positive affect (e.g. feeling happy) tend to perceive more losses than agents in whom a positive affective state was not induced (Isen, Pratkanis and Slovic, 1984).

Using a similar roulette game task, Isen and Gueva (1987) found that positive-affect subjects require a high probability of winning in order to accept a risky bet (vis-à-vis the control group). This empirical result also indicates that the experience of positive affect shape risk perception and risky preferences in a way that individuals become quite averse to highly risky prospects.[157]

3.1.2 Emotions shape preference for risky prospects

A growing number of decision theorists including psychologists and economists have examined the impact of particular emotions among other affective processes and states on risk judgments, preferences, and decisions. Lerner and Keltner (2001) investigated the role of fear and anger in guiding actual risky choice. They drew on insights from emotion appraisal theory[158] to hypothesize that an agent's perception of an uncertain task or situation with little individual control elicits fear, whereas her or his perception of a decision task involving control and certainty will trigger anger in agents.[159] Lerner and Keltner's experiments indicate that fear shapes individual preference for a sure-thing option against risky prospects. In this case the experience of fear may be associated with risk-aversion. The elicitation of anger feelings, on the other hand, prompts individuals to opt for a risky gamble (i.e. to reveal risk seeking preferences).[160]

---

157 Isen, Nygren and Ashby (1988) designed a gambling task that in which some roulette players were induced to experience a positive feeling state. Subjects bet less on gambles with a high probability of losing. The phenomenon was explained by showing that positive affective states frame risk preference so that individuals become more averse to losses.

158 According to Roseman and Smith (2001, p.3), the essence of appraisal theory is the claim that a particular emotion is 'caused' by (not necessarily conscious and deliberate) evaluation of the agent's relationship with the environment.

159 For details, see Lerner and Keltner (2001, p.147).

160 Lerner and Keltner also investigate whether (and how) anger and happiness shape risk perception and preference. Their experiments indicate that individuals under the experience of happy and angry feelings often underestimate risk associated with available prospects.

3.2 Affective states and processes guide probability judgment and (utility) value ascription

Recall that the EU framework is built on the notion of subjective probability. This interpretation of probability does not imply a commitment to the view that emotions and feelings guide the decision machinery by virtue of the influence they can exert on agent's probability reasoning. Probability is called 'subjective' in so far as it reveals an individual's beliefs in the occurrence of certain events (with expected utility outcomes).[161]

In this section, I challenge the basic expected utility framework by offering evidence that affective factors guide probability and utility judgments in very specific ways. For instance, on perceiving a risky task an agent may experience certain emotions that steer her or his inferences in relation to probabilities and outcomes of salient options. Based on experimental work, Johnson and Tversky (1983) showed that individuals in whom negative feeling states were induced tended to make more pessimistic probability judgments. [162]

Other studies have examined the connections between positive feeling states and probability inferences. Nygren, Isen, and Dulin (1996) induced a positive affective state (feeling of happiness) in some of their subjects, and then asked all participants to estimate the probability of outcomes associated with three lotteries and to reveal their betting preferences. Nygren and colleagues found that positive affect subjects made more optimistic judgments of the prospects of winning (than control subjects). Nevertheless, positive affect subjects bet lower amounts of money in lotteries that involved large monetary losses and therefore seemed to be more averse to risk than other participants of the experiment. These results show that the experience of positive affect give rise to inferences that in turn reveal some "cautious optimism". The standard EU approach

161 Ramsey (1931), Savage (1954).

162 In this experiment, a group of students is asked to read newspaper articles about deaths of students referring to potential causes of death, such as disease (leukemia), fire, floods and murder. Another group read happy newspaper articles. Then all estimated the probability of dying from different diseases, hazards or other types of violence. They found that individuals under the experience of negative affective states drew more pessimistic judgments.

failed to accommodate this phenomenon. A risk-as-feeling account can deal with this apparent puzzle to the extent that it incorporates the assumption that happy individuals are risk averse because they fear a shift to a negative affective condition and relaxes the various SEU assumptions that make reference to risk judgment and perception independent from acts and events.[163]

3.3 The explanatory gain of a 'risk-as-feeling' account

The 'risk as feeling' account succeeds in accommodating patterns of behavior that remain anomalies for EU. This empirical account seems to employ a re-isolation strategy of theorizing (i.e. re-isolative move) to the extent that it replaces the hypothesis of expected utility maximization with the empirical hypothesis that risky choice is shaped by feeling states in important ways. The risk-as-feelings account replaces some assumptions that constrain the predictive and explanatory capacities of economic analysis of risky choice because, such as (SV_6) and (SV_7). By describing the ways in which affective states shape risk judgments and decisions, the risk as feelings account can deal with economically significant anomalies, such as overweighting small probabilities and preference reversal. [164] Unlike some amended versions of the expected utility model that that rely on non-linear probability functions to incorporate risky choice anomalies, the 'risk-as-feeling' account allows for a clearer picture of decision-making under risk/uncertainty since it uncovers some affect-driven processes causally relevant roles to the occurrence of manifest behavior.

---

163 According to Isen, Nygren and Ashby (1988) ,
"Persons who are in a positive affective state and are considering the positive outcomes in risky situations may focus, for decision-making, on the probability of winning; the impact of affect on consideration of these outcomes may be most evident on probability estimation or weighting. When making decisions under uncertainty, these individuals may place less importance on the actual value of a positive outcome and more importance on the likelihood of occurrence. On the other hand, when considering possible losses, these persons may focus on how the loss will feel (its subjective utility) rather than on its likelihood; the deciding factor in people's behavior with regard to losses may be the impact of the affective state on utility estimation" (p.716).

164 The 'risk-as-feeling' account also offers an alternative interpretation of a famous expected utility anomaly, simultaneous gambling and insurance behavior, by referring to the agent's experience of emotions and feelings such as dread and excitement.

The 'risk-as-feeling' perspective offers a schematic explanation of why agents see a negative relationship between risk and benefit (despite the positive correlation in nature). It suggests that agent's feelings about hazards shape agent's risk perception and judgments about probabilities and utilities associated with risky prospects (Alkahami and Slovic, 1994).[165] Based on the risk-as-feeling account, we can predict that individuals will be more willing to buy insurance against terrorism than a policy covering their home being affected by floods if a terrorist attack taking place at the time of decision evokes emotionally vivid images of these prospects with negative consequences.[166] The risk-as-feeling approach predicts that some risky judgments and choices are made independently of probability variations.[167] The EU framework cannot accommodate such explananda items. The 'risk-as-feeling' account, in its turn, deals with apparently puzzling instances of behavior by suggesting that individuals under the experience of certain affective factors do not change their perception of risk and judgments of emotionally salient hazards (e.g. nuclear power, or exposure to toxic chemicals) even when informed of their low probabilities of occurrence (Slovic, Finucane, Peters and MacGregor, 2002a).

The 'risk-as-feeling' perspective carries some revolutionary implications. It relaxes the assumption that probabilities and outcomes ascribed to risky prospects are independent (i.e. orthogonal). More importantly, it uncovers some cognitive and emotional processes

---

165 Similarly, Peters and Slovic argue: "when the affective evaluation was favorable, the activity being judged was seen as having high benefit and low risk; when the evaluation was unfavorable, risks tended to be seen as high and benefits as low" (1996, p.1429).

166 Perhaps this behavior pattern does not apply to individuals whose homes have been damaged by floods.

167 Rottenstreich and Hsee (1999) studied a behavior pattern that the EU framework cannot accommodate. In one experimental setting, participants were asked to state the highest sum of money they would be willing to pay to avoid an unpleasant situation that occurred with different levels of probability. The task was to choose between a money loss of $20 (an emotionally pallid outcome) or a brief painful electric shock (an emotionally salient or vivid outcome). They found out that decisions were very sensitive to probability changes when subjects perceived an emotionally neutral outcome at stake. The dollar-value participants placed on an unknown prospect was $1 for (p=0.01) and $18 (for p=0.99). However, observed choices responded very little to probability changes when the outcome was associated with strong emotionally salient outcomes. The value ascribed to a prospect shifted from $7 to $10 when probability varied from p=0.01 to p=0.99. In another experiment Rottenstreich and Hsee students had to state the amount of money they were willing to pay to receive a positive outcome. One positive outcome was a $500 discount off university tuition (a pallid outcome); another was a $500 coupon for a dream vacation (emotionally salient outcome). They found that participants were sensitive to probability variations for pallid outcomes (tuition discount) whereas emotionally vivid outcomes prompted choices insensitive to probability changes.

with major roles in (causal) production of manifest behavior.[168] The risk as feelings account encounters some sources of objection. The account is more properly characterized as an empirical hypothesis in need of further development so as to represent a theoretical model of choice that is as general as the expected utility approach and able to effectively describe the chain of affective and cognitive events and processes with major roles in risky judgment and decision.

4. The Affect Heuristic Model

Based on further empirical research, Paul Slovic and his collaborators sketched an account of risky choice behavior called affect heuristic model. The latter purports to unveil mental processes with crucial roles in production of actually observed patterns of behavior (Slovic, Lichtenstein, Reid and Coombs, 1978). This account of risky choice differs more significantly from the EU framework to the extent that it does not resort to an explanation of choice behavior in terms of utility axioms and probability rules. The heuristic model dispenses with the isolated set of *explanantia* elements underlying Savage's subjective expected utility approach. Rather, the affect heuristic model offers an explanation of behavior in terms of cue-based mental shortcuts (heuristics, inferential procedures) that give rise to risk judgments and decisions. The distinctive trait of a heuristic approach is that it relaxes the criterion of rational choice behavior as expected utility maximization and points to a perspective on bounded rationality (Simon 1997, Gigerenzer and Selten 2002). The affect heuristic model can be also regarded a re-isolative move that regards emotions and feelings as significant driving forces behind agent's detection of favorable choice alternatives, judgments about salient risky prospects and the consequences associated with them, and selection of a favorable course of action that meets her or his perceived level of aspiration.

168 Loewenstein, Weber, Hsee and Welch suggest that the 'risk as feelings' account allows for predictive and explanatory improvements. To them, "addition of feelings… makes risky choice more predictable, both within and across different decision domains and contexts" (2001, p. 274)

4.1 What is an affect heuristic like?

Affect heuristics amount to mental representations of potential gains and losses attached to positive or negative feeling states, which in turn guide inferential and decision processes significant for risky choice behavior. [169] Accordingly, feeling states offer credible cue-based heuristics for quick and (often effective) probability judgments and value estimation by boundedly rational agents. This heuristic model hypothesizes that the thinking mind comprises two systems: the affective-experiential and the logical reason-algorithmic systems.

The affective-experiential system enables an individual to automatically process information on objects in the world regarded as significant. The affective system operates by means of body-related cues from earlier experiences (and mental images associated with them). The logical reason system refers to a slower, effortful mode of information processing that enables agents to undertake higher-order cognitive activities (e.g. means-end reasoning, Bayesian updating) relevant to the occurrence of choice behavior.

4.2 How the affect heuristic model deals with risky choice anomalies

The affect heuristic approach yields understanding of explananda phenomena that failed to be accommodated by the EU framework. It explains how actual individuals estimate an inverse relationship between risks and benefits (Slovic, Finucane, Peters, and MacGregor 2002a, b). It offers understanding of why risky stocks (prospects) are expected to offer low benefits or positive consequences (Ganzach 2001). It predicts that financial analysts relying on affect heuristics to make judgments on unfamiliar stocks will associate low risks with high benefits. The affect heuristic approach can accommodate the well-known EU anomaly called preference reversal. According to Slovic and Weber (2002), these

169 As Finucane, Alkahami, Slovic and Johnson (2000) put the issue,

"We propose that people use an affect heuristic to make judgments. That is, representations of objects and events in people's minds are tagged to varying degrees with affect. People consult or refer to an affective pool (containing all the positive and negative tags associated with the representations consciously or unconsciously) in the process of making judgments. Just as imaginability, memorability, and similarity serve as cues for many probability judgments (e.g. availability and representativeness heuristics), affect may serve as a cue for many important judgments. Using an overall, readily available affective impression can be far easier - more efficient – than weighing the pros and cons or retrieving from memory many relevant examples, especially when the required judgment or decision is complex or mental resources are limited. This characterization of a mental shortcut leads us to label the use of affect a 'heuristic'. (p.3)

puzzling instances of choice behaviors occur because individuals have difficulty in estimating positive outcomes that a gamble may yield. The affect heuristic model offers an ingenious way of resolving some preference reversals by suggesting that addition of a small loss to a gamble enhances its attractiveness in comparison to a sure gain. The affect heuristic model relies on the explanatory statement that an individual's perception of a possible loss elicits affective experiences and emotionally vivid images that in turn guide cue- based inferences in relation to the stakes and their affective (hedonic) outcomes.[170]

Furthermore, the affect heuristic model offers a straightforward account of biased probability judgments that called the expected utility framework under question. It predicts that individuals recall emotionally vivid events or situations more easily than emotionally pallid ones.[171] For instance, Slovic and colleagues take that agents' reliance on affect heuristics as explaining why they sometimes overestimated the probability of highly publicized causes of death (e.g. cancer, AIDS, or murder) while at other times underestimating the probability of dying from less publicized causes (diabetes, asthma and tuberculosis, for example).[172]

4.3 Explanatory and predictive capabilities

Clearly then, the affect heuristic model expands the scope of explanatory tasks pursued by models of risky choice. In this respect, the model fits the pursuit of an account that yields explanatory progress as scope expansion. I shall now proceed to explain this. The affect heuristic model accommodates new explananda items like EU anomalies as well as other patterns of behavior already covered by standard economic analysis. In addition, the affect heuristic model may contribute to explanatory progress as enhanced causal

170 Slovic and his collaborators subjected this argument to empirical testing. In one experiment, half of the participants were asked to choose between a sure gain of $2 and a lottery of the type ($9, 7/36; $0, 29/36). The other half were asked to decide between the $2 and ($9, 7/36; $-5c, 29/36). The result was that preferences did not shift. Only 33.3% preferred a gamble without a monetary loss to a sure gain, whereas 60.8% of subjects opted for the lottery (when its attractiveness was enhanced by adding a loss outcome).
171 In this respect, the affect heuristic perspective yields an alternative explanation of probability judgment to one based on purely cognitive heuristics such as 'availability' (Kahneman and Tversky 1982).
172 The underlying idea is that the heavily publicized risks of death are represented in the individual's mind as emotionally arousing events. Less publicized causes of death often evoke pallid affective states and therefore may be recalled less vividly.

articulation since it describes some mental processes with major roles in risk perception, probability judgment, value ascription, and therefore production of manifest behavior.

Recall that the affect heuristic model does not assume that risk judgment and decision are driven only by cognitively demanding mental processes but also by emotional states and processes. Unlike some behavioral models of choice that accommodate emotional factors to the explanantia set, the affect heuristic model does not presuppose that emotional processes and feeling states always produce distorted outcomes. Rather, it suggests that under particular circumstances the agent's experience of some moderate emotions will enable her or him to identify an important risky task and behave adaptively to it. [173]

The affect heuristic model yields explanatory progress for risky choice behavior compared with the conventional expected utility model, firstly by providing an account of risky choice behavior in terms of (causally relevant) mental processes and thus adding to our understanding of how risky choice behavior happens in the world. Secondly, the affect heuristic approach paves the way for an explanation of human decision-making machinery that uncovers the roles played by emotions, feelings, and higher order cognition in decision-making under risk/uncertainty. Nonetheless, the affect heuristic model is limited in its ability to track the causal chain of events that put together risk judgments and preferences; emotion and higher order cognition; and emotional factors and manifest choice behavior. In an attempt to circumvent this limitation, I apply a two-level model of decision-making behavior (suggested in chapter 4) to the domain of risky choice.

5. A two-level explanation of decision-making under risk

This section shows that an account of behavior at the complementary levels of mind and brain yields some explanatory advancement to the extent that it (a) sheds further light on the significance of described causal processes, (b) offers a detailed description of the decision machinery that gives intelligibility to the complex phenomena of risky

173 See Loewenstein et al 2001, p.271

inferences and choices, and (c) promotes understanding of why emotions are non-negligible ingredients of a (causal) explanation of risky choice behavior.

5.1. Emotion, risk perception and detection of a choice task

A two-level description of the decision machinery shows that emotions like fear and its variants guide individual's risk perception in significant ways. At the perception of an ecologically significant danger, dread reactions prompt an individual to mobilize her sensory and conceptual systems (physically realized by neural pathways involving sensory thalamus, sensory neocortex and amygdala among other brain structures) so as to draw inferences about the decision problem at hand and how to respond adaptively to it. The explanatory account presupposes that basic emotions like fear guide the individual's capacity to identify important choice tasks and to find a solution to the frame problem. This is at least partly so because activation of particular emotions prompt cue-based inferences about previously experienced situations and attentional processes, which in turn contribute to mental representation of an urgent task. This account of the first stage of the decision machinery with two levels of analysis shows that dread and other affective reactions are significant and robust processes with major roles in agent's capacity of risk perception and detection of an important choice task.

There seem to be two neural pathways involved in human ability to process information in a potentially risky situation.[174] One is the thalamus- amygdala route and another is the thalamus-sensory cortex-amygdala trajectory (LeDoux 1996/1998). Both information processing pathways shape risk perception by activating the dread component of risk and what has been called the 'unknown-risk factor'. The former component of risk is mobilized quite automatically, without access to conscious awareness and control. The latter component, unknown risk, is activated in a voluntary and controlled manner with

174 Dread is a variant of fear emotion that is elicited by one's perception of a very dangerous situation or event. There is evidence that the amygdala plays a crucial role in mobilizing and processing information about fear-eliciting tasks and as such may also be part of the physical brain substrate for dread reactions to a perceived risk. The underlying idea is that the amygdala mobilizes processing activities that enable an individual to draw quick inferences about the nature and significance of an uncertain task. Our two-level account of risky decision-making takes dread reactions as informational outputs from the detection stage of the decision machinery. They may guide subsequent risk judgment and decision-making through the mobilization and regulation of specialized cognitive and bodily processes relevant to individuals' selective search for satisfactory prospects.

room for deliberation. Provided that the amygdala and prefrontal cortical structures comprise the physical underpinnings of the agent's capability for immediate perception of a risky problem situation and are also involved with emotion processing, development of a two-level of risky choice behavior is worthwhile since it sheds light on the non-negligible roles emotions play in mobilizing and coordinating activities whereby the agent can exercise her or his capacity to detect an urgent choice task.[175]

5.2 Emotions and selective search for satisfactory (risky) prospects

Emotions and feelings offer cues as to possible choice alternatives. In so doing they serve the functional role of regulating mental activities that enable individuals to engage in selective search for risky alternatives.[176] Emotions are thought to play a role in guiding processes relevant to the working of the decision machinery. This is because emotions accompany hedonic states that evoke vivid mental images of good or bad previously experienced situations or events. Such pieces of information stored in memory and learning systems enable individuals to draw quick inferences about available risky prospects that yield positive consequences. From this perspective, emotions may serve the functional role of activating simple (cue-based) mental shortcuts that seek favorable risky prospects quickly.

Note that one advantage of a two-level account of risky choice is that it draws on findings from neuroscientific research that inform economists among other decision theorists that particular mental processes (with brain underpinnings) play significant roles in selective search for choice alternatives. The model suggests that specialized brain structures physically realize information processing activities whereby selective search happens. The amygdala, the prefrontal cortex, the hippocampus and the basal ganglia constitute the physical substrate for the agent's mental capacity to mobilize and regulate attention,

175 For instance, an individual's detection of an important threat or danger may activate a subcortical pathway producing dread reactions to a perceived risk. When a risky situation is perceived as an unknown or unobservable distribution of possible outcomes, it may activate higher-order cognitive processes by which the agent can draw careful (though slower) inferences about the nature and significance of the problem situation.

176 Emotions are interpreted as behavioral reactions that follow bodily changes and perception of a danger or opportunity posed to the agent, while feelings refer to conscious experiences of emotions. See details in chapter 4.

learning and memory processes by which agents seek favorable courses of action and their consequences. A two-level account of risky choice behavior reveals that the described neural structures and information processing activities play significant roles for production of manifest behavior. Some of these brain structures are particularly involved with emotion processing (e.g. the prefrontal cortex and amygdala). One positive implication of a two-level explanation of decision-making is that offers a detailed account of the complex machinery for risky choice that gives greater intelligibility to the complex phenomenon of risky choice. In addition the two-level account sheds extra light on the role emotions play in guiding cognitive activities by which individual can exercise her or his capacity to find satisfactory choice alternatives.

5.3 Feelings and the selection stage

The proposed two-level account of risky choice behavior presupposes that feeling states (i.e. conscious experiences of emotions) shape the inner operation of the decision machinery in a quite systematic (non-random) manner. It incorporates the explanatory statement that feeling states are motivational factors with significant roles in production of manifest behavior. From such perspective, feelings mobilize inferential and evaluative processes by which the agent exercises her or his capacity to assess salient options and to pick out one that meets perceived aspirations. Following the neuroscientist's vision that feeling states guide individual's judgments about hedonic outcomes associated with unknown prospects, the model suggests that positive feelings serve as cue-based sources of inferences about rewarding prospects (that ought to be pursued). Negative feeling states yield reliable cues about punitive risky prospects (that ought to be avoided).

As we can see, the two-level model allows for a detailed explanation of the decision machinery that offers scientific understanding to the complex phenomenon of risky choice behavior. The model adds to the development of a behavioral account in terms of affect heuristics since it clarifies why feelings are crucial elements of the last stage of the decision machinery (i.e. they carry information about vivid images relating to previous choice tasks and their consequences). Furthermore, the suggested model suggests that the individual can exercise her or his mental capacity to opt for a favorable prospect because

she or he is equipped with specialized brain structures and activities causally sufficient for determination of manifest behavior. At this final stage of the decision machinery, the prefrontal cortex, amygdala and nucleus accumbens comprise a network processing activities that operate in parallel so as to enable the agent to assess some salient choice alternatives and to select a favorable one. Given that the abovementioned structures are particularly involved with emotion processing, it can be shown that a two-level account of behavior throws extra light on the major role feelings play in guiding agent's selection of a choice option that meets her or his aspiration level.

To summarize, a two-level model is worth developing since it offers a detailed account of the decision machinery that extra intelligibility to the complex phenomenon of decision-making under risk/uncertainty and its puzzles. This is because it uncovers those mental states, events and processes (physically realized by particular brain structures) that are significant and robust to observed decision outcomes (i.e. manifest behavior). Additionally, a two-level model contributes to an improved explanation of choice behavior patterns to the extent that it exposes causally significant ways in which emotions and feelings shape decision-making under risk. The model implies the view that emotions activate and mobilize specialized processes (also describable as information-processing activities) whereby agents identify an urgent risky choice task, coordinate limited computational resources to seek alternative risky prospects, and select a salient course of action that satisfies her or his perceived aspiration level. The suggested account may add to the explanation of decision-making under risk to the extent that it specifies those mental processes and states (physically realized by particular brain structures) playing major roles in production of actual behaviors. In so doing, a two-level approach might contribute to the development of an explanatory theory that promotes progress as enhanced causal articulation and scope expansion.

6. Concluding remarks

Decision-making under risk/uncertainty is a complex phenomenon that cannot be satisfactorily covered by the standard expected utility framework. In this chapter, I examine some developments from the standard expected utility model in terms of de-

isolative and re-isolative moves in behavioral theorizing aimed at increased predictive and explanatory power. The underlying idea is to show that emotions play non-negligible roles in production of manifest behavior and therefore ought to be incorporated in (explanatory) economic accounts of risky choice behavior.

Based on some variants of the expected utility framework, I suggest that some behavioral models that incorporate some expected emotions (e.g. regret, rejoice, disappointment) so as to predict novel facts and behavior patterns that remain anomalies (e.g. preference reversals, overweighting small probabilities, simultaneous gambling and insurance behaviors) can be interpreted as de-isolative moves. Many behavioral models of choice under risk relaxed some debatable SEU assumptions like independence so as to incorporate previously excluded explanatory items that are now thought to play major roles in actual judgment and decision-making behavior. Some accounts supplement the previously isolated set of explainers with the ideas that expected emotions (e.g. regret and disappointment) shape actual individuals' inferences and choice in non-negligible and non-random ways. For instance, regret theory incorporates the idea that one's comparison of the outcome of an action with a different action in the same state of the world may elicit the emotion of regret that in turn changes the way the expected utility function is calculated. Quite similarly, disappointment theory relaxes the assumption of independence (i.e. the sure thing principle) and supplements the expected utility model's set of explanantia with the assumption that risky tasks require the agent to make comparisons of outcomes of the same action in different states of nature that may trigger the emotion of disappointment that shape selected outcomes (section 2). The first conclusion of this chapter is that these refined versions of the expected utility can deal with behavior patterns covered by the standard expected utility model and predict novel facts (especially phenomena regarded as anomalies). Then, regret theory and disappointment theory can be interpreted as examples of de-isolative moves that have contributed to the economist's pursuit of explanatory progress in terms of scope expansion. Yet their silence about those states, processes or mechanisms with major roles in production of actual behavior seems to constrain the possibility of achieving progress as improved causal articulation.

The current work advances the idea that behavioral researchers in search of reformed models of decision under risk may profit from developing models that draw on a process-description strategy. With this in mind, I scrutinized two behavioral accounts that deviate from the incremental strategy of theorizing. The risk-as-feelings account can be taken as an early case of re-isolative move that challenges the vision that some explaining items isolated by the utility framework like the independence assumption are necessary for an explanation of manifest behavior and its puzzles. Alternatively, the risk-as-feelings account relies on experimental studies to describe non-negligible roles emotions and feelings play in risk perception, judgment and decision-making (section 3). The second concluding remark of this chapter is that the risk-as-feelings account strategy of replacement allows for improved predictions and explanations of some risky choice anomalies (e.g. preference reversals, common consequence effect, loss aversion, simultaneous gambling and insurance). It is important to stress, nevertheless, that the risk as feelings account cannot be properly regarded as an analytical alternative to the expected utility model. Rather it resembles an empirically grounded treatment of puzzling patterns of risky choice behavior.

This chapter implies the view that the affect heuristic model is an empirical model that adds to the risk-as-feelings account. It purports to be the first step towards a mechanistic explanation of decision-making under risk. The affect heuristic model replaces the hypothesis of expected utility maximization with some affective processes by which risk judgment and decision-making happens (section 4). My third conclusion is that the affect heuristic model contributes somehow to progress as improved causal articulation to the extent that it describes affective processes and states by which behavior patterns are causally produced. The affect heuristic model allows for a clearer picture of decision-making under risk that better predicts and explains expected utility anomalies. Yet it is still far from being a broad account that traces the complex chain of emotional and higher order cognitive processes and states that produce economically important behaviors. This is at least partly so because it does not offer detailed description of the roles emotional

and cognitive processes play in mobilization and guidance of a non-random causal chain of events that produce risky choice in the real world.

In order to add to the development of an explanatory account of risky choice behavior on lines with the affect heuristic model, this chapter suggests a two-level model of decision-making under risk that yields a quite rich description of the decision machinery. Its improvement has to do with the capacity to expose a network of those emotional and cognitive processes (physically realized by brain structures with capacity for information processing activities) with causally significant and systematic roles in production of patterns of risky choice behavior. Then, a two level account yields extra intelligibility to the controversial phenomenon of decision under risk. In so doing, a two-level account of risky choice behavior is worth developing to the extent that it offers a clearer picture of the links between emotion, higher order cognition, inferences and choice in a complex network of causal dependencies. This may contribute to progress as causal articulation and also scope expansion.

I end this chapter with three remarks. One is that we must re-build the mode of explanation characteristic of standard economic analysis if we wish to develop genuinely explanatory accounts of behavior. Until now there is not a behavioral theory of decision under risk that is superior to the expected utility framework in terms of generality, simplicity and congruence with reality. In the current work, I put forth a two-level account of risky choice behavior that yields a strategy to undertake such an important task. Its distinctive trait is to incorporate a physical brain of analysis to maintain that emotion among higher order cognition are significant components of the machinery of choices in a boundedly-rational world full of uncertainties about alternative choice prospects and their consequences.

CHAPTER 7
EMOTION AND EXPLANATION OF PROSOCIAL ACTION

> Selection may favor distrusting those who perform altruistic acts without the emotional basis of generosity or guilt because the altruistic tendencies of such individuals may be less reliable in the future.
>
> R. Trivers, *"The Evolution of Reciprocal Altruism"*

> One reason why the human decree may take cultural roots is that the design of the brain tends to facilitate its practice. It is likely that the simplest forms of some behaviors necessary for realizing the human decree, such as reciprocal altruism and censure, is merely waiting to be awakened by social experience. We have to work hard at formulating and perfecting the human decree but to some extent our brains are wired to cooperate with others in the process of making the decree possible. This is the good news. The bad news of course, is that many negative social emotions, along with their explanation in modern cultures, make the human decree difficult to implement and improve.
>
> A. Damasio, *Looking for Spinoza*

1. Introduction

It is a fact that human beings show a strong disposition to cooperate with others. Many contribute to public goods and charities, such as orphanages, mental hospitals, and museums, in some cases anonymously and in large groups.[177] Some even risk their own lives to save strangers from danger. These instances of cooperative behavior cannot be adequately predicted or explained by the conventional model of rational agency, centered on the hypothesis of self-interest maximization. Rather they point to the apparent puzzle of human prosociality, i.e., the human animal's strong willingness to cooperate with

177 Anonymous blood donation is another type of genuine altruistic behavior that cannot be accommodated by the self-interested agent model.

others and to punish cheaters even at great personal cost (Bowles and Gintis 2003a, b and Gintis 2003). In this chapter, I examine some models of prosocial (cooperative) behavior that have been designed in reaction to decision researchers' doubts about whether the standard economic account of social behavior did not remove some explanatory factors (e.g. prosocial preferences, norms and emotions) that play major roles in production of the phenomena under study.

Similarly to the two previous chapters, the current work puts forth the idea that some behavioral models of cooperative choice that aim at improved explanations (predictions) of prosocial anomalies can be regarded as de-isolative moves, whereas some others that uncover those mental processes or mechanisms for choice behavior make re-isolative moves.

The present chapter has modest purposes. It makes three main contributions. One is to show that incremental analysis and process- description approach are two reformist modeling strategies employed by behavioral researchers that pursue, respectively, two goals that mean improved explanatory performance: scope expansion and enhanced causal articulation. Another is to advance the argument that certain emotions like anger, shame and guilt ought to be incorporated in an improved theoretical representation of cooperative behavior since they play key (rather than negligible) roles in mobilization and regulation of a complex system of information-processing by which real-world agents exercise their mental capability for effective prosocial (antisocial) judgment and decision-making. The third contribution is to suggest that a behavioral model of cooperative behavior that incorporates a mental and a neural level of analysis is worth developing since it provides rather detailed explanation of how emotional and higher order cognitive processes participate in a chain of events that produce manifest prosocial (cooperative) behavior and therefore contribute to the attainment of the goals of scope expansion, improved causal articulation and enhanced causal penetration. The underlying idea is that a two-level account does not only specify how emotional and higher order cognitive processes participate in the (causal) chain of events that produce patterns of prosocial (and antisocial) behavior but also digs into the neural foundations (i.e.

specialized brain structures and processing activities) that execute physically those major states and processes whereby manifest behavior comes about.

The line of argument developed in this chapter proceeds as follows. Section 2 discusses the standard self-interest model proposed by social scientists and raises doubts about whether it removed some explanatory items that play major roles in production of various instances of cooperative behavior. It goes on to suggest that systematic prosocial anomalies indicate that the self-interest model failed to account economically relevant patterns of behavior and they prompt economists to employ reformist modeling strategies, often regarded as de-isolative moves. Section 3 analyzes some behavioral models that incorporate explanatory factors like social preferences, norms of fairness, norms of public contributions and moral emotions so as to attain the goals of scope expansion and causal articulation that mean improved explanatory performance. It argues that the behavioral model of strong reciprocity pursues the ideal of causal articulation since it offers an account of prosocial behavior in terms of its driving forces.[178] Section 4 proposes a two-level model of cooperative choice that purports to be a development of accounts that attempt to specify the major processes or mechanisms for manifest behavior. It goes on to claim that an account of cooperative choice behavior at the mental and neural levels has an improved explanatory performance, since it offers a detailed explanation of how emotional and higher order cognitive processes (physically realized by particular brain structures and processing activities) bring about actual prosocial acts (even in large groups where the possibility of future interaction is remote). Section 5 summarizes the overall argument and draws conclusions.

2. The self-interest framework

Largely due to a long-standing tradition, economists often equate the notion of rational agency with individual's pursuit of her or his self-interested preferences. Likewise, the economist Francis Edgeworth draw on the assumption that the first principle of

178The discussion is informed by Gintis (2000), Henrich and Boyd (2001), Fehr and Gächter (2000, 2002) and Fischbacher *et al* (2001).

economics is that man is only actuated by self-interest (1881, p.16).[179] It seems that many economists of the 19th and 20th centuries regarded the assumption that individuals only pursue their own self-interest as a domain (of applicability) assumption. The latter serves to remove some explanatory items like altruistic preferences, norms of fairness and prosocial emotions that are thought to play non-negligible effects on actual people's behavior that need to be absent for a well-behaved model of rational choice to apply.

More recently, George Stigler (1975, 1986) rhetorically claims that self-interest is the granite on which the palace of economics analysis was erected. According to him, the idealizing assumption that individual choices are self-interested may serve as a negligibility assumption. It excludes explanatory factors (ethical considerations among other normative or affective commitments, for instance) that are thought to play negligibly small effects on behavior regularities observed within the market society, such as child rearing, mating and religious behavior.[180] Stigler claims that the self-interest model is a useful simplification of psychology of choice that can be applied to a wide range of phenomena and therefore contributes to the economist's pursuit of the goal of scope expansion.

The above theoretical representation of human action suggests that standard explanation of economic choice behavior in terms of the principle of self-interest maximization is now regarded quite questionable. The economist Amartya Sen (1990) faults the self-interest model for simplifying individual preferences and choices, and thus failing to predict behavior patterns deviating from its predictions, such as ethically committed choices, donations to charities, and so forth. He complains that the self-interest model offers a quite distorted theoretical representation of the reality of choice behavior and therefore failed to account for economically relevant anomalies among other new phenomena (e.g. why many people prefer to buy goods from socially responsible companies even at higher prices and make anonymous contributions to charities). To Sen,

179 See chapter 2 for details on the history of economic accounts of choice behavior.
180 He even argues: Let me predict the outcome of the systematic and comprehensive testing of behavior in situations where self-interest and ethical values… are in conflict. Much of the time, most of the time in fact, the self-interest theory… will win (Stigler 1986, p.176)

the assumption that behavior is solely self-seeking is a too strong simplification of agent's motivational architecture implying a commitment to a controversial psychological doctrine of psychological egoism.[181]

More recently, behavioral researchers (Ernst Fehr, Samuel Bowles, Herbert Gintis, Nancy Eisenberg, Dan Batson, for instance) suggest that the motive for any voluntary act is not necessarily the possibility of giving some benefit for the actor. In response to that, economists, psychologists and even biologists develop amended versions of the self-interest model that deal with puzzling (though significant) patterns of prosocial behavior.[182]

2.1 The enlightened self-interest model

The early reformed versions of the self-interest model are developed in response to decision researchers' questions about whether the basic model of human agency did not exclude factors like other regarding preferences that play non-negligible roles in production of important patterns of behavior and therefore restricted the explanatory performance of standard economic analysis.

The enlightened self interest approach seems to result from an incremental modeling strategy that is here viewed as a de-isolative move. Versions of the enlightened self interest model seems to supplement the set of explainers of the basic self-interest model (e.g. individuals have rational preferences represented by a utility function, they make optimal judgments about other's social expectations and behavioral dispositions and make choices consistent with the principle of rationality as self-interest maximization) with new explaining items like "individuals long term preferences depend on other's preferences and well-being". As a result, the enlightened self-interest model can account for a larger set of explanandum phenomena. The enlightened self-interest model can deal with behavior patterns that remained unexplained by the standard self-interest model,

181 Psychological egoism can be defined as 'the theory that all our ultimate desires are self-directed" (Sloan and Wilson 1998, p.201)

182 For a very interesting account of how biologists explain cooperative (altruistic) behavior, see Sloan and Wilson (1998) and Gintis et al (2005).

such as cooperation within the family and money contributions to charities. The above phenomena are explained by reference to the hypothesis that a rational individual selects a cooperative course of action if this offers a means to satisfy her own long-term preferences. The enlightened self interest approach predicts the occurrence of prosocial acts when they yield a larger payoff to the agent than those associated with any alternative course of action.

Many amended self-interest models are based on Becker's (1976) model of altruistic behavior. They account for cooperative behavior in terms of an individual's willingness to meet her or his long-term utility (whose arguments include family members' consumption vectors).[183] Note that these reformed models make interesting de-isolative moves to the extent that they accommodate explanatory elements (e.g. possibility of improving one's own condition by assisting others) without questioning the conventional outcome (payoff)-oriented treatment of choice.

Some variants of the enlightened self interest model can account for an apparently puzzling phenomenon - individuals make contributions to public goods against their own self-interest. This is because these models take into consideration (previously excluded) explanatory factors like improved self-image and reputational gains playing significant effects on cooperative behavior. For instance, James Andreoni (1988, 1989) offers a version of the enlightened self-interest model that incorporated the explanatory variable of 'warm glow effect'. He suggests that prosocial behavior prompts in the agent positive affective experiences that inform her or him about positive outcomes associated with contributions to a public good (despite the material cost). In light of this model, individuals are led to cooperate with others because a prosocial strategy pays off.
Also driven by an incremental reformist modeling strategy, Landes and Posner's (1978) offer a refined version of the self-interest model that includes the factor that cooperative behavior yields positive reputational gains. The model can deal with new explanandum phenomena, such as helping behavior and philanthropic acts (money contribution to

183 For details, see Zamagni (1995).

charities).[184] Glazer and Konrad (1996) develop an enlightened self-interest model that includes the assumption that the agent contributes to joint activities with the aim to signal wealth and obtain reputational gains associated with cooperative behavior. In so doing the model can effectively predict apparently puzzling phenomena like named philanthropy and donations to charities.

As we can see, these refined versions of the self-interest model can accommodate explananda items that failed to be predicted by the standard self-regarding preference framework. This is at least partly so because they include some other-regarding considerations and preferences with non-negligible roles in determination of prosocial acts. The above models remain within the framework of self-interest framework to the extent that prosocial behavior is explained as the best possible strategy to meet the benefactor's long-term preferences.[185] If this is so, these de-isolative moves contribute somehow to attainment of the goal of scope expansion.

2.2 Limitations of the enlightened self-interest model

Enlightened self-interest models can deal with various instances of prosocial behavior, such as voting in national elections, individual contributions to charities and public goods and named benevolence. Yet they cannot explain (or predict) patterns of cooperation that occur on an anonymous basis. Nor can they deal with situations in which the individual cooperates with others even to the detriment of her/his own interests.

Additionally, the enlightened self interest model cannot satisfactorily explain a relevant instance of cooperative behavior called 'psychological altruism'. The latter indicates that some patterns of prosocial behavior are driven by agent's concern with another's goals, preferences and values regardless of her or his material self-interest.[186] This is at least partly so because enlightened self-interest models also seem to have removed some items

184 The authors offer no details about the sources of interdependencies across individual utility functions.
185 Glazer and Konrad show that most contributions to universities, philharmonic orchestras, and museums are not anonymous. Of a total of 2,240 individual monetary contributions to Pittsburgh Philharmonic Orchestra in 1991, only 29 were anonymous. Alumni made 1950 individual donations to Yale Law School, but only 4 were anonymous. In 1989 and 1990, Carnegie-Mellon University received 5,462 money contributions but only 0.3% of the donors did not reveal their names.
186 According to Sober and Wilson, psychological altruism indicates that individuals care about others for their own sakes (1998, p.6)

with major roles in real-world thinking and decision-making (and therefore also isolated too much or wrongly). According to Zamagni (1995), there is systematic evidence that reputational gains or enhanced self-image are not sufficient motives for important patterns of prosocial behavior, such as anonymous donations and helping strangers who cannot reciprocate in the future.

3. Emotions and norms in an explanation of prosocial behavior

Economists acknowledge that the standard model of self-interested preferences failed to yield adequate explanations (predictions) of actual cooperative behavior (e.g. Frey 1992, Fehr and Schmidt 1999 and 2000). In response to that, they work hard to offer models of choice that include new explanatory items with significant roles in production of actual behavior. This section examines three reformed models of prosocial choice that incorporate explanatory factors - emotions and norms - constituting non-negligible driving forces behind patterns of cooperative behavior: (a) the commitment model; (b) the norm of fairness model (c) the norm-emotion (social preference) model, and (d) the altruistic punishment model. I go on to show the positive implications they bring to the achievement of scope expansion and causal articulation.

3.1 Robert Frank's commitment model

In his *Passions within Reason,* Frank doubts whether the basic self-interest model succeeded in accounting for economically significant patterns of cooperative behavior, such as cooperation in one-shot PD games and hard-core altruism.[187] Following Hirschleifer's (1987) ideas that emotion and feelings offer credible guarantees of one's readiness to cooperate (and to retaliate in case of defection), Frank suggested a commitment model that explicitly takes emotions as driving forces behind empirical regularities of economic behavior and therefore ought to be incorporated in the model's set of explaining items.

Frank's model relies on the hypothesis that emotions involve facial and bodily expressions that serve as cue-based heuristics, by which individuals resolve commitment

187 One example of hard-core or genuine altruism is anonymous blood donations.

problems.[188] The intuitive idea is that human emotional expressions signal an agent's readiness to cooperate (even to the detriment of her or his material self-interest). Provided that human emotions cannot be experienced and expressed in a perfectly deliberate and voluntary fashion, they offer a 'wired in' mechanism, whereby individuals can effectively coordinate their expectations concerning prosocial behavior strategies.

The commitment model draws upon ideas of evolutionary biology (and its perspective on individual selection). It assumes that individual disposition to react emotionally to the perception of others' prosocial (and antisocial) actions is a behavioral trait that yields a differential in reproductive success.

Frank sees emotions and feelings as conferring material advantages to individuals faced with the social commitment problem. To explain this, he takes a large population containing both cooperators and cheaters playing a PD game. Prosocial emotions (e.g. sympathy, shame, guilt) experienced by potential cooperators motivate them to cooperate with others despite personal costs.[189] Defectors in their turn do whatever is required to pursue their personal material self-interest.

Frank further asks why (and under what conditions) individuals with the ability to behave emotionally do better than others. He finds that the facial and bodily expressions that accompany emotions serve as signals of the agent's readiness to enjoy the benefits of mutual cooperation and incur the cost of retaliating defectors. This communicative (signaling) function of emotions facilitates selective pairing among cooperators. When this happens, cooperators cooperate with cooperators and punish free-riders. Based on the evolutionary rule, cooperators have greater reproductive success than agents that cannot

188 Some examples of the commitment problem are participating in a risky business joint venture, getting married and rearing children.
189 Bowles and Gintis (2003b) argue that prosocial emotions amount to mental and physiological reactions that prompt individuals to decide over cooperative acts (even at a personal material cost).

enjoy the gains of mutual cooperation. As a result, the fraction of individuals equipped with the prosocial behavioral trait tends to increase in the population.[190]

The commitment model suggests that emotional factors do the explaining (at least partly) of why the agent may opt for a course of action undermining her or his material self-interest. From such perspective, cooperative acts are explained by reference to the emotional processes and states that trigger in human agents an urge for cooperation with cooperators and retaliation of cheaters. Putting somewhat different, the model is the first attempt to sketch a proximate mechanism for cooperative behavior patterns that do not square very well with the conventional approach to choice behavior. By pointing to the fitness consequences of emotion-driven (choice) strategies, the commitment model also offers a description of an ultimate mechanism that explains the evolution of cooperation between biologically unrelated individuals (in large groups). [191]

To summarize, the contribution of the commitment model is to incorporate explicitly new explanantia items (e.g. emotions and feelings) with major roles in production of behavior. Accordingly, it deals with explanandum phenomena that remain puzzles (e.g. genuine altruism, cooperation among cooperators in large groups of unrelated agents, punishment of shirkers or free riders) as well as new phenomena associated with the human capability for prosociality (e.g. agents engage in cooperative acts like getting married and some contributions to public goods persist over time). The improved explanatory power of the commitment model (vis-à-vis the self-interest model) lies on its capacity to describe some mental processes by which manifest prosocial behavior come about in the real world.

3.2 Norms of fairness and social preferences as non-negligible explanatory factors

Based on systematic experimental studies, behavioral researchers found that social norms and social preferences play significant roles in production of apparently puzzling behaviors with economic significance and therefore ought to be incorporated into an

190 Note that the commitment model shares similarities with the biological approach to prosocial behavior. It differs from the model of reciprocal altruism to the extent that it relaxes the assumption that prosocial behavior requires agents to coordinate their expectations of future interaction and reciprocation.

191 For detailed and thoughtful analyses of Frank's commitment model, see Vromen (1995), Elster (1998), and Verbeek (2002).

improved theoretical representation of cooperative choice behavior. Some economists and psychologists found compelling evidence that, unlike predictions made by the basic self-interest model, actual people reject positive sums of money even to the detriment of their material self-interest. Güth, Schmittberger and Schwarz's (1982) explain rejections of positive amounts of money in their ultimatum games by suggesting that individuals follow norms of fairness that motivate them to punish greedy offerers even to the detriment of their material self-interest.[192]

Based on a two-staged ultimatum game, Binmore, Shaked and Sutton (1985) confirmed the hypothesis that internalized norms of fairness shape the agent's decision over punishing social exploitation by rejection of positive money sums even at a material cost. [193] In this case, norms of fairness prompt rational subjects to coordinate their expectations on a fair division of the "cake". Public game experiments also offer evidence that norms of fairness guide individual preferences for prosocial behavior in quite robust manners. Fehr and Schmidt (1999) found that many individuals driven by norms of fairness contributed to public good investments even when the material payoff of free-riding is quite large.[194] Frey and Weck-Hanneman (1984) account for the degree of the agent's compliance with formal and informal contract in terms of her or his judgments about

192 In a well-known experiment, participants were divided into proposers and respondents. The experimenters instructed the proposer to suggest a division of money ($10) between herself or himself and an anonymous recipient. If the respondent rejected the offer, both parties would receive nothing. The possible allocations varied from $9.50 to the proposer and $0.50 to recipient to an even split ($5 to offerer and $5 to respondent). The standard model of rational economic man predicts that the offerer will keep $9.50 for herself (himself) because this is the amount that maximizes her (his) material self-interest. The recipient, in turn, will accept any positive amount, since something is better than nothing. However, actual behavior violates these predictions. Proposers tended to offer around 50% of the total sum, whereas respondents rejected positive sums of less than 30% of the amount.

193 In the first stage of the experiment participants were asked to divide $20 with an anonymous student as in the dictator game (recipient cannot reject offer) and 76% of the 161 students opted for an even share-out. All 161 were then asked if they would pay $1 to punish a cheater and reward a fair allocation. In this case, 74% were more willing to pay the cost of punishment and divide $10 with a fair proponent than share $12 with an unfair individual.

194 Based on a meta-study of 12 public good experiments, Fehr and Schmidt found that agents often contributed from 40% to 60% of their endowment. Repeated interactions among individuals did not increase cooperation rates. Instead, contribution rates declined in the final rounds of public goods experiments. Almost 73% of participants contributed nothing in the final stages. Fehr et al suggest that cooperators lower their contribution levels on seeing others free riding and violating norms of fairness.

unfair allocations. Then, a worker's perception that his employer deviates from norms of fairness might produce shirking and/or low morale (Greenberg 1990).[195]

As we can see, a behavioral account of cooperative behavior that incorporates norms of fairness as non-negligible explanatory items can deal with some apparent puzzles of human prosociality, such as individual rejections of money in ultimatum games, contribution to public goods even at a material cost and compliance with particular formal and informal contracts, etc. In this case there are grounds for suggesting that models that take social norms seriously might contribute to improved explanatory performance of economic analysis of human cooperation in the real world.

3.3 Prosocial emotions and strategic cooperative behavior

The psychologists P. Straub and K. Murninghan (1995) seem to suggest that norms of fairness are not sufficient to better explain why people reject positive money offers in ultimatum games. They hypothesize that some emotions like anger and wounded pride play a non-negligible role in that and therefore ought to be taken into consideration by analytical models. Based on a series of ultimatum game experiments, they offered a schematic explanation of why and how significant rejections of money in ultimatum games occur. First, the respondent perceives an unfair offer. Second, judgments that a norm of fairness is violated activate anger and wounded-pride reactions. Third, the negative affective states associated with social emotions (anger and pride) serves as motives for rejecting monetary offers. [196, 197]

Many other studies offer evidence that emotions are major elements in production of patterns of prosocial (cooperative) behavior. Bosman and van Winden (1999) designed a "power-to-take" experimental game that investigated whether individuals' perceptions of those responsible for violations of a norm of fairness and experienced prosocial emotions

195 More recently, Fehr and Schmidt (2000) found that norms of fairness were the rationale for wage premium policies.

196 In a similar line of reasoning, Pillutla and Murnighan (1996, p.222) argue that experience of anger feelings partly does the explaining of rejection of monetary offers and moralistic aggression.

197 It seems that the experience of some emotions like anger responses (along with other prosocial or moral emotions) is dependent (at least partly) on the agent's ability to attribute causality and identify violator(s) of an internalized norm.

shaped the observed intensity of moralistic aggression (i.e. punishment of individuals who do not reciprocate cooperation with cooperators). [198, 199] They found that the participants' experiences of emotions led them to exhibit the following pattern of behavior: the higher the take rate (i.e. an evidence of social exploitation and violations of fairness), the more intense were the activated feelings of irritation, contempt and envy. These emotional reactions and feeling states prompt them to punish the take authority.[200]

Note that an account of prosocial behavior that incorporates social norms and emotions does not only provide an explanation of why cooperate with cooperators and punish cheaters at a personal cost in a very abstract sense. It might be very useful to accommodate prosocial choice anomalies with economic relevance. For instance, it can offer an interesting treatment of the phenomenon of 'fair monopolistic pricing'. Although economists often ridicule the talk of "fair prices", many are embarrassed at the perception that the basic self-interest framework failed to deal with price setting behavior for desirable goods such as tickets for Rolling Stones concerts or World Cup soccer games. The standard self-interest model predicts that suppliers would exploit situations of excess demand by raising prices. However, this is not what really happens. This apparent paradox is resolved once we take emotions and norms of fairness as non-negligible explanatory factors. Of course it is unlikely that suppliers' pricing decisions is only explained by norms of fairness and equity. Nevertheless, their expectations that their customers may take their price increases as violations of norms serve to deter monopolistic exploitation. Therefore, 'fair prices' may be a means of avoiding the consumer's emotional (behavior) response to an unfair or morally outrageous price.[201]

As we can see, these extended versions of the self-interest approach can deal with anomalies as well as behavior patterns already covered by the conventional account of social behavior. This means that they are conducive to the goal of scope expansion that

198 It is also known as a type of dictator game.
199 Experimenters divided 78 undergraduate students into pairs: a "take authority" and a "responder". The take authority decides which fraction of the responder's income is to be transferred to her or him. Responders who feel the take authority is unfair may retaliate (punish) authority by destroying income.
200 Bosman and van Winden noticed that take authorities chose high rates (mean 58.5, median 66.7 and mode 70.0). 21% of the 39 responders destroyed some or all income, and 6 destroyed it all.
201 See Friedman (2001)

means improved explanatory capacity. Nonetheless, the available accounts of cooperation do not offer a clear treatment of how emotion and higher order cognition participate in a causal chain of events that produce manifest prosocial behavior and its apparent puzzles. In this case, we are still in need of developing alternative models that pursue the goals of scope expansion as well as of causal articulation or penetration. In what follows, I analyze two behavioral accounts that seem to make re-isolative moves so as to fill the abovementioned blank. They are interpreted as developments of the self-interest model that replace the conventional rational choice explanations in terms of self-interested preference maximization and optimal beliefs about one's own and other's behavioral dispositions and choice strategies with accounts that describe emotion-driven heuristics and mental processes like altruistic punishment and rewarding.

3.4 Behavioral models that uncover processes or mechanisms for cooperative behavior

Some contemporary behavioral economists like Ernst Fehr, Simon Gächter among others emphasize that the self-interested preference framework isolated wrongly the major factors that do the explaining of systematic patterns of human cooperation and therefore ought to be revised. Based on recent experimental studies, they sketch an explanatory model of prosocial behavior that unveils (causally) relevant emotional and higher order cognitive processes and states. Their account is built on the hypothesis that individuals are ready to cooperate with cooperators and to punish cheaters even to the detriment of their material self-interest.

Based on a series of public goods experiments, Fehr and Gächter (2000, 2002) investigate their working hypothesis of strong reciprocity among individuals during social interaction. Their aim was to uncover the driving forces behind agent's decisions over prosocial acts. Fehr and Gächter devised three different experimental conditions. In the first, subjects interacted with the same individuals in ten rounds of "personal (partner) treatment". For another ten rounds, they were randomly allocated to new groups after each round ("stranger treatment"). Thirdly, for the "perfect stranger treatment", subjects were reassigned to a new group after each round and informed that they would not interact with any other subject more than once.

Fehr and Gächter found that individuals were willing to punish cheaters even if they had to pay a price for doing so. In the personal (partner) treatment, the possibility of punishment led individuals to increase their individual contributions over time and most participants made 100% contributions in the last round of the game, despite the anonymity conditions of the experimental setting. In the stranger treatment, the possibility of punishing shirkers prevented cooperation declining. When punishment was not possible, contribution rates declined over time and subjects tended to contribute nothing in the last round. The results are summarized in the chart below:

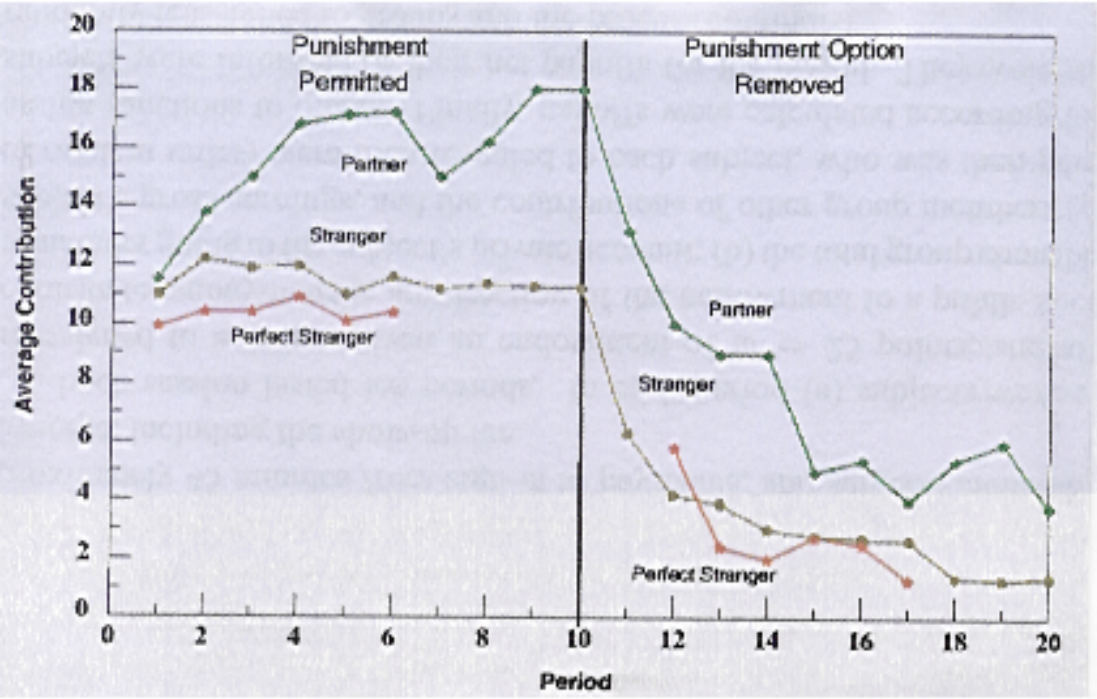

Figure 8.1:
Average contributions when punishment of cheaters is "possible" and "not possible"; when individuals interact with the same individuals during the 10 periods of the game ("partner treatment"), in the "stranger" and "perfect stranger" situations. Source: Bowles and Gintis (2002, p.7)

The abovementioned empirical findings led Fehr and Gächter to outline two mental processes by which patterns of prosocial behavior among unrelated agents in large groups may arise and persist over time – altruistic rewarding and altruistic punishment. Altruistic rewarding prompts the agent to be ready to cooperate with cooperators, whereas altruistic

punishment motivates the agent to punish cheaters or free-riders (even at the expense of one's perceived material self-interest).[202]

Fehr and Gächter (2002) designed another experiment to investigate the robustness of altruistic punishment to cooperative behavior in large groups of individuals (interacting anonymously). Students were asked to play a public goods game with real monetary stakes in two different situations: with and without the possibility of punishing free riders. Subjects were divided into groups of 4 members and each was given 20 monetary units (henceforth MUs). The money could be directed to a public account or not (i.e., subjects were free to contribute between 0 and 20 MUs to a public good). For every amount invested in the project, each member of the group would receive 0.4 MU (the group return would be 1.6 MUs) regardless of her or his own contribution. The cost of investing in the public project was 1 MU.[203] If nobody contributed, every agent would receive only 20 MUs. If they all contributed to the public good, each could receive 32 MUs. Given that the public investment decisions are made simultaneously and interactions occur on an anonymous basis, Fehr and Gächter study what happens when punishing free-riders/ cheaters is possible.

In the punishment condition, individuals could punish each of the other group members after being informed about their contributions. The decision over punishment would be to assign between 0 and 10 points to the cheater. The cost of punishing to the punishing member was 1 MU and to the punished agent was 3 MUs for every assigned point. The punishment decisions were also simultaneously made. The game was repeated for six periods and participants learned that they never met more than once in a group. The results were telling: by the last part of the experiment, 40% were making maximum contributions to the public good, i.e. 20 MU, and almost 80% contributed more than 15MU when punishment was allowed and actually imposed on shirkers. On the other

202 Note that altruistic punishment (and rewarding) relates to the notion of evolutionary altruism, since the agent increases the average payoff of the group at the expense of her or his own reproductive fitness (material resources)

203 Note that, if the return of the investment were lower than the cost (0.4 MU < 1 MU) contributing nothing would be the optimal self-interested strategy.

hand, 75.6% of subjects contributed less than 5MU, and 60% of them did not contribute at all when there was no punishment opportunity.

The above described experiments found three empirical regularities of behavior. Firstly, subjects making above-average contributions tended to punish participants making below average contributions (free riders). Secondly, the lower the rates of contributions to public goods, the stronger the feelings of anger experienced by cooperators. Thirdly, the impact of punishment on contributions was immediate at the turning points between "with punishment" and "without punishment" situations. One interesting result of Fehr and Gintis' work is that it uncovers the non-negligible roles some moral emotions like anger and fear play in mobilization of mechanisms for prosociality (i.e. altruistic punishment and cooperation). In this case, emotions can be regarded as (non-negligible) proximate causes of cooperative behavior in large groups of people when the prospect of repeated interaction is remote.

On similar lines, Herbert Gintis (2000) provided a strong reciprocity model that aims to explain why (and how) prosocial behavior emerges and persists. His analytical treatment incorporates the presumption that prosocial emotions combine with higher order cognitive processes to resolve the puzzle of cooperative behavior.[204] In order to test his hypothesis, Gintis designed a public good experiment in which a group of n individuals can contribute an amount c> 0 to a public account and receive a benefit, b. The expected fitness of contributing to a public investment (given that other parties also contribute) is denoted by π. The only Nash equilibrium amounts to universal defection when agents interact non-repeatedly. However, a cooperative equilibrium outcome can occur when some conditions are met (e.g. the game is continued). Therefore, the agent's expected payoff can be written as follows:

$$\pi = (b - c) + \delta \pi, \qquad (1)$$

δ is the probability of the game continuing in the subsequent period and gives a fitness value of π. With some algebraic manipulation,

204 See also Bowles and Gintis (2003a)

$$\pi = (b\text{-}c).\ (1+ \delta + \delta^2 + \delta^3 + \ldots) = \frac{b - c}{1 - \delta} \qquad (2)$$

An agent contributes to a public good, if $(b\text{-}c) / (1\text{-}\delta) \geq b$. This condition is needed because those who don't contribute to public goods receive b in the current period and nothing afterwards. By rearranging the terms, cooperation evolves in a repeated public goods game setting, if and only if,

$$c/b \leq \delta \qquad (3)$$

The above inequality shows that cooperation will flourish in a model with self-interested actors if the probability of future interaction (denoted by δ) is high. Yet Gintis emphasizes that individuals, who are strong reciprocators, tend to cooperate with cooperators and to punish cheaters even when (δ) is low. This apparent puzzle is resolved by adopting an evolutionary perspective. Groups of individuals within a population often risk extinction, so a group with a high proportion of strong reciprocators may perform better (in material or reproductive fitness terms) than one with many self-interested members. Gintis suggests that the behavioral trait of strong reciprocity increase its frequency in the overall population and consequently society can enjoy the fruits of mutual cooperation. Following the frequency dependent (evolutionary) logic, Gintis can explain why and how cooperation among between individuals in large groups arises and persists over time.

My own view is that Gintis' strong reciprocity model supplements Frank's and Fehr's treatment of puzzles of human prosociality. The three models regard emotions as domain-specific mental algorithms (cognitive activities) with major roles in guiding judgments and decisions. Despite the different emphasis the abovementioned authors give to the signaling function of emotional expressions, they take prosocial emotions as powerful triggers of the altruistic punishment and rewarding mechanisms (and therefore are causally relevant to the occurrence of cooperative behavior).

In short, Frank's commitment model, Fehr and Gächter's model of altruistic punishment and Gintis' account of strong reciprocity includes emotions as non-negligible explanatory

items and this partly explains why they can deal with apparently puzzling behavior patterns, such as cooperation in one-shot PD games and persistence of cooperative behavior within large groups where reciprocation is unlikely. These models seem to be re-isolative moves since they replace the standard accounts of prosocial behavior in terms of rational preferences and beliefs consistent with the principle of self-interest maximization with an explanation by reference to affect-driven heuristics and mental (emotional and cognitive) processes that are relevant to the occurrence of manifest choice behavior. One positive implication of such theoretical developments is that they contribute to explanatory progress as scope expansion. For example Fehr's and Gintis' models offer an improved understanding of why contributions to public goods may decline over time. This may happen when cooperators (strong reciprocators) react emotionally to violations of norms of contribution and reduce their own contributions to punish free-riders. [205, 206]

Although the above behavioral models of cooperation resolve some economically relevant anomalies, they did not offer a detailed account of how emotion, higher order cognition and manifest prosocial choice behavior are related in a complex network of causal (inter) dependencies.[207] In an attempt to circumvent this problem and pursue the goals of scope expansion and causal articulation, I shall provide an outline explanation of the complex phenomenon of human cooperation at two levels of analysis. The idea is to show that a detailed account of the decision machinery is necessary to track the major roles emotion and higher-order cognition play in (causal) production of the apparent puzzles of human cooperative behavior.

---

205 See also Fischbacher, Gächter and Fehr (2001).

206 Evidence for altruistic punishment is reported in Dawes, Orbell and van der Krag 1986, Camerer and Thaler (1995), Roth (1995), Fehr and Gächter 2000, McElreath et al (2001/2003) and Binmore (1995). Fehr and Schmidt (1999) offer evidence that confirms the empirical validity of 'altruistic rewarding '.

207 According to Bowles and Gintis,

> Whereas we think evidence is strong that prosocial emotions account for important forms of human cooperation, there is no universally accepted model of how emotions combine with more cognitive processes to affect behavior (2003b. p. 433).

4. In search of a two-level explanation of human cooperation

As in the previous two chapters, this section is an attempt to apply a two-level model of decision-making to the phenomenon of prosocial behavior. It suggests that an explanation of cooperation that includes a brain level of analysis is worthwhile to the extent that it allows for a description of those mental states and processes (physically realized by specialized neural structures and processing activities) with central roles in (causal) production of manifest behavior. The model shows that a model of prosocial choice at the mind-brain levels has the merit of exposing the causal significance and robustness of affective processes for cooperative behavior. In addition it offers grounds for the claim that some emotions (e.g. anger, fear, shame, guilt, sympathy) are key components of the decision machinery by which cooperative acts arise in the real world. Emotions enable agents to exercise their mental capacity for undertaking information processing activities that enable them to detect an important problem situation involving interpersonal conflicts, to mobilize computational resources to encounter satisfactory choice alternatives and to select a prospect that satisfies their perceived (felt) aspirations and concerns.

4.1 The roles of emotions in the detection stage

In our perspective, individuals are endowed with particular sensory and conceptual systems (also describable in terms of specialized physical brain structures and information processing trajectories) that enable them to exercise their ability to identify important hazards or opportunities posed by particular social environments (Cosmides and Tooby 2000). For instance, fear and anger may inform the agent about the importance of a detected commitment problem. Perception of another's violation of a norm of cooperation will automatically evoke anger reactions that in turn mobilize the individual's attentional processes to respond adaptively (e.g. by punishing free-riders or reducing contributions to a public good).

Following Adolph's and Damasio's systematic studies of the neural underpinnings of cooperative behavior, the present model suggests that three main brain structures allow agents to exercise their capacity to identify quickly a commitment problem and to

respond adaptively to it. Recall from Chapter 4 that the sensory thalamus, the sensory cortex and the amygdala perform activities involved in perceptual representations (inferences) of an environmental stimulus that constitute the detection stage of the decision machinery. The outputs from these inferences become informational inputs and elicit brain areas responsible for emotion processing. The merit of an account that includes this neural level of analysis is that it points to the centrality of emotions in shaping those decision problems arising from situations that involve interpersonal conflicts of interest. In addition, a description of the first stage of the decision machinery that specifies those neural foundations offers empirical evidence that emotions are non-negligible cues for immediate detection of an important task. Damage to brain areas like amygdala and sensory cortex impairs the agent's mental capacity to draw recognize emotions in others' faces, to draw quick and effective inferences about their behavioral inclinations and to assess the significance of a particular social interaction at hand (Adolphs, Damasio, Tranel, Cooper and Damasio 1999, Ruby and Decety 2001, Damasio 2003).

4.2 Prosocial emotions and the selective search stage

On the basis of meticulous experimental work, Adolphs (2003) suggests that there is a second set of neural structures, on which the human ability to reveal prosocial behavior depends. It includes the amygdala, striatum and orbitofrontal cortex. Chapter 4 also suggested that these brain regions are responsible for mobilization and regulation of those information processing activities whereby the agent can find alternative ways of resolving the perceived choice task (problem situation). One advantage of a description of the decision-making mechanism at the physical level is it shows the important roles some emotions like shame (guilt) and anger play in guiding the stage in which individuals mobilize their computational resources to find alternative courses of prosocial (antisocial) action. They may activate attentional, learning and memory systems (physically executed by the amygdala and hippocampus among other brain structures) through which the agent can exercise her or his mental capabilities for encountering satisfactory choice alternatives or behavioral strategies. At this processing stage, emotions might be useful to inform the agent's whether she or he (or a third party) deviated from a social norm (e.g.

making a fair contribution to public investments). They prompt cue-based inferential activities by which boundedly rational agents find possible solutions to the problem of social cooperation that at least satisfy her perceived goal priorities or aspirations. In this case, a two-level account of the decision machinery is worth developing since it contributes to causal articulation within the mental realm (i.e. it specifies the chain of emotional and higher order cognitive events that prompt agents to search find alternative courses of action).

4.3 Emotions and selection of a course of prosocial action

A two level account of prosocial choice presupposes that certain brain structures are also responsible for the agent's mental capacity to pick out a course of (prosocial or antisocial) action. Based on Adolphs (2001) and Damasio (2003), it can be suggested that the left prefrontal, medial, anterior and posterior cingulate cortices regulate processing activities associated with executive functions such as monitoring, planning and decision-making; by working together they enable individuals to construct 'mental models' of social environments including representations of other people, inferences concerning their social reasoning strategies and behavioral inclinations, and evaluations as to how to coordinate one's own expectations and interests with those of others with whom one interacts. The above-mentioned higher order cognitive processing activities are performed by cortical regions that keep various connections with areas like the amygdala and the nucleus accumbens. These two structures play a central role in this final stage of processing activities, since they constitute the systems of motivation and motor behavior. According to Damasio and his colleagues, the prefrontal cortex, the amygdala and the nucleus accumbens constitute the driving forces behind the agent's mental capacity to select a salient choice alternative that satisfies (or surpasses) her or his perceived (felt) aspirations, priorities, concerns or values.

A neural level of description of decision-making gives some understanding of why the abovementioned brain structures comprise a system linking sensory information relating to a certain social stimulus with inferences concerning the affective experiences or

motivational values associated with it. In so doing, a physical brain level of analysis yields an improved understanding of the crucial roles feeling states (i.e. conscious experiences of emotions) play in determination of cooperative behavior even in large groups of unrelated people facing little chance of future reciprocation.

It is important to stress that emotions shape the third processing of the decision making system in a systematic fashion. They are associated with positive or negative affective states, i.e. feelings that offer cues to rewarding or punishing outcomes to be pursued or avoided. In this sense, emotional experiences or feelings serve as somatic markers or affective cues (Damasio 1994) that individuals can use to make quick and often adaptive decisions. For instance, the experience of anger feelings is associated with negative hedonic factors that prompt individuals to preserve the future benefits of mutual cooperation by punishing cheaters despite the costs to themselves. Put in another way, anger may mobilize moralistic aggression mechanisms such as altruistic punishment that partly explain cooperative patterns even in large groups. Moreover, anticipation of anger in others gives rise to quick inferences concerning the long-term costs of cheating that lead one to select a prosocial course of action. Accordingly, sympathy, compassionate love and gratitude are accompanied by specialized feeling states in the actor that trigger a strong desire to reciprocate cooperation despite the high material payoff associated with cheating or free-riding (Trivers 1971; Fessler and Haley 2001).

Guilt and shame are two other important emotions that play non-negligible roles in an explanatory account of cooperative behavior. Guilt is elicited when the agent realizes her/his deviation from a social norm has caused some harm to another. This prosocial emotion mobilizes the agent's computational resources to draw quick inferences concerning the nature or extent of damage caused by its cheating. Guilt feelings involve negative hedonic factors that motivate the agent to pursue a course of action repairing or compensating for the social damage inflicted.[208, 209, 210] Shame is triggered when an

208 In this case, the human ability to anticipate guilt feelings also serves as a deterrent for cheating in so far as it points to its negative hedonic outcomes (vis-à-vis the benefits of mutual cooperation).

individual perceives that she or he has violated a socially approved norm of conduct. But unlike guilt, activation of shame depends on whether members of a group or community are aware of the violation and ready to retaliate. Shame plays an important role in inducing cooperation in a potential cheater due to its capacity to mobilize specialized cognitive processes that guide her or his inferences in relation to the negative consequences associated with being punished and excluded from the gains of mutual cooperation.[211] Then, guilt and shame may guide individuals' computational resources (attention, memory, and learning processes) to favor alternatives complying with socially accepted and internalized norms and therefore allow them to enjoy the (large) payoffs of prosocial strategies.[212]

In this case, a description of the decision machinery that also incorporates a neural level of analysis seems to be worthwhile since it informs us that emotions may be significant and robust processes for mobilization and propagation of prosocial patterns of behavior. It highlights that particular brain lesions destroy the agent's emotion-triggering capabilities, and more importantly impair the individual's capacity to relate certain features of a social situation to the feeling states that guide cue-based selection of a prosocial or antisocial output (Adolphs 2001 and Damasio 2003).

4.4 Neuroscientific evidence on the significance and robustness of emotions

James Rilling and his collaborators (2002) adopted magnetic resonance techniques to investigate whether particular brain structures constitute the physical foundations of individual's capacity to mobilize emotional and higher order cognitive processes by

---

209 Thus the expectation of feeling guilt may prompt one to forego the temptation of free-riding.

210 Experiments based on iterated PD games confirm the hypothesis that inducing guilt feelings in agents enhances their readiness to cooperate. Iterated ultimatum game experiments indicate that proponents feeling guilty after making unfair or greedy offers will later tend to behave in a prosocial manner. See Fessler and Haley (2001/2003)

211 Bowles and Gintis (2002) designed a public goods experiment to show how internalized shame and guilt guided cooperative behaviors. They found that prosociality evolved in large groups because punishment of cheaters evinced the community's disapproval of norm violation, which led free riders to feel shame and guilt. Such negative affective experiences caused cheaters to make more contributions to the public account.

212See Gibbard (1992) and Bowles and Gintis (2002) for insightful analyses of the connections between emotions and norm-following behavior

which actual cooperative behavior comes about in the real world. Based on two iterated PD game experiments, Rilling and his colleagues scanned the brains of 36 women players. One experiment attempted to isolate the brain correlates of cooperation and non-cooperation from all other factors. They found that the pattern of cooperation changed when participants were told that they would play the game with another woman or with a computer. Another experiment focused on the differences between human and computer interaction. Rilling and colleagues found that mutual cooperation was the most common outcome and pairs of participants adopting a CC strategy in the first round often followed a cooperative strategy in the subsequent round.[213] However, the frequency of cooperation decreased in the final rounds of the game and individuals cooperated less when dealing with a computer.[214] Their empirical results confirm the hypothesis that emotions mobilize and regulate mental processes by which mutual cooperation emerges and is sustained. To Rilling and his collaborators, prosocial behavior is dependent on the activation of brain regions involved with rewarding and positive emotion processing, such as the nucleus accumbens, the anterior cingulate cortex and the ventromedial frontal/ orbitofrontal cortex. The neuroscientists found that agents' selection of a CC strategy was systematically correlated with activation of the striatum; orbitofrontal cortex, and anterior cingulate cortex (i.e. brain structures that constitute the physical substrate for the human ability to behave prosocially). Moreover, the experience of positive feelings associated with cooperation seems to reinforce subject's preference for a prosocial act. The strategy of mutual cooperation is also positively associated with activation of the somatosensory cortex (a brain structure responsible for the ability to detect a social task and motivate mutual cooperation (Rilling *et al.* 2002, p.401). The CD and DC outcomes, in their turn, involve deactivation of striatal neurons (the striatum is a portion of the brain associated with rewarding a prosocial act). Subjects of the experiment found that the defection strategy in response to social exploitation was selected for self-preservation and punishing cheaters. Subjects selecting a DD strategy experienced guilt feelings that prompted them to shift to a cooperative outcome.

213 CC is the mutually cooperative strategy adopted by players A and B. There are other three possible outcomes. Player A cooperates and player B defects (CD); player A defects and player B cooperates (DC), and both players defect (DD).
214 Andreoni (1995) explained this issue by referring to cooperators reacting angrily to perceived social exploitation.

As we can see, an empirical study of prosocial choice that also takes a physical brain level of description into account gives additional confirmation to the hypothesis that emotions (e.g. shame, guilt, anger and sympathy) play central roles in coordination of some mental processes causally relevant to the occurrence of manifest behavior in a non-random manner. Certain prosocial emotions serve evolutionary cognitive functions that enable agents to exercise their capacity (a) to detect a task that involves some interpersonal conflict, (b) to search of potentially satisfactory alternatives, and (c) to select of an option that satisfies or exceeds their own or others' perceived aspirations and goal priorities. This is partly so because these emotions have specific physical brain underpinnings. A physical level of analysis offers a way of examining whether (and why) certain physical brain structures (e.g. amygdala and prefrontal cortex), playing major roles in execution of the agent's capability for decisions under risk, are also important to activation and regulation of activities that bring about prosocial emotions and emotion-based cooperative acts.[215]

Furthermore, the strength of a two-level account of cooperative behavior lies on its capacity to cover a wide range of prosocial phenomena (including apparent puzzles), such as helping family members and friends, assisting strangers from danger, named benevolent behavior, anonymous contributions to public goods and charities, cooperative behavior in large groups of biologically unrelated people where the chance of future reciprocation is remote, etc. Remember that other behavioral economic models are often tailor-made to resolve particular prosocial anomalies or puzzles. By adding a new category of explanatory factors viz. emotions, a two-level model of decision-making behavior also offers understanding of how different domains of economic behavior – intertemporal choice behavior, decision under risk/uncertainty and cooperative behavior might be related in a complex economic world of causal (inter)dependencies. In this sense, it might give behavioral scientists the chance to pursue the scientific goal of progress as unification (see Mäki 2001d).

215 According to Antonio Damasio, some brain structures (e.g. amygdala and prefrontal cortex) significant for the agent's capacity to feel social emotions (e.g. anger, shame, guilt, sympathy) also regulate information processing by which human beings can deploy their repertoire of social emotions in order to assess the benefits of prosocial or antisocial action and to quickly select a satisfactory option (2003, p.155).

An explanation of behavior that includes a neural level of analysis is worth developing since it informs economists among other behavioral researchers that emotions are central ingredients of the decision-making system that evolved to deal with economically relevant commitment problems. More importantly, I argue that an account of prosocial choice behavior that incorporates a physical brain level of analysis is worthwhile since it exposes those mental processes (physically realized by specialized neural structures and information processing activities) that bring about and sustain economically important patterns of behavior. Putting somewhat different, addition of a brain level of description offers a detailed account of the decision machinery for prosocial behavior that deals explicitly with the complex issue of causation in economics. In my own view, it is worth pursuing the abovementioned task so as to attain the goal of causal articulation and penetration. A two-level account helps us to track the places of mental events like emotion and higher order cognition in the causal structure of the economic world.

5. Concluding remarks

Casual observation and experimental research tell us that human beings are inclined to cooperate with others and punish cheaters or free-riders even at great cost to themselves. This chapter has examined models of choice that were developed to explain (predict) the apparent puzzle of human prosociality. Given the complexity of the 'explanandum' chosen, I began by discussing the standard account of the self-interest model and the theoretical refinements appended therein to deal with behavior patterns that remained anomalous (section 2). I have argued that the self-interest model and its amended versions can explain important patterns of cooperation, such as named contributions to public goods and charities. Yet I showed that the enlightened self-interest model and its variants are de-isolative moves that still leave out explanatory factors playing major causal roles in production and maintenance of prosocial behavior, and therefore failed to explain empirical regularities like types of anonymous cooperation, without any gains in terms of social reputation or chance of future reciprocity.

With this in mind, I analyzed behavioral models that incorporate two (previously excluded) explanatory items that are significant for the occurrence of actual prosocial

behavior - emotions and norms (section 3). Based on insights from biology, Frank's commitment model was interpreted as the first re-isolative move that replaces the assumption of rational choice as maximization of one's self-interested preferences with the assumption that emotions lead to rational outcomes undermining the agent's perceived self-interest. One of the advantages of the commitment model is that it can explain (predict) behavior patterns that remained unexplained by the conventional model (e.g. why people cooperate even in one-shot PD games, why people reject positive money offers and contribute to public goods) and also covers new phenomenon (i.e. it offers an account of why and how such patterns of cooperative behavior persist over time). In so doing, the commitment model contributes to pursuit of the goal of scope expansion. By describing emotions and norms that seem to be proximate causes for prosocial behavior, Frank's model also pursues the goal of improved causal articulation. This leads me to my first concluding remark. Frank's commitment model provided the first step towards the development of a genuine explanatory account of cooperation (altruism) since it evinced the role of emotions and feelings in an explanation of prosocial patterns that are observed in various patterns of behavior with economic relevance. I also discussed some behavioral models of choice that make de-isolative moves by incorporating explicitly social preferences and norms of fairness in the preference framework. I showed that models that take emotions and norms as non-negligible explanatory factors might be conducive to scope expansion since they can deal with behavior patterns that remained puzzles or anomalies to the basic self-interested preference approach (e.g. large contributions to public goods and altruistic behavior). However, these variants of the enlightened self-interest model cannot offer a straightforward account of why real-world individuals adhere to norms of cooperation even at great cost to themselves. For that reason, I went on to examine two behavioral models of choice that uncover emotion (and cognitive) driven processes or mechanisms by which actual prosocial behavior arises and persists over time. I suggested that Fehr and Gächter's accounts of prosocial behavior in terms of altruistic punishment and Gintis' model of strong reciprocity are re-isolative moves that relax the core assumption of substantive rationality as utility maximizing behavior and replace it with the idea that actual agents rely on cue-based heuristics to make quick and adaptive choices in situations involving social interaction (and potential

interpersonal conflicts of interest). I argued that Gintis's strong reciprocity model is an ingenious treatment of human prosociality to the extent that it exposes driving forces behind the emergence and maintenance of actual human cooperation in large groups. This leads me to the second conclusion. Together with the commitment model, the strong reciprocity model and accounts of behavior in terms of altruistic punishment/rewarding give a clearer picture of those processes that are causally relevant to actual behavior and its apparent puzzles. This may be partly so because each model seems to offer a partial (though complementary) explanation of the various ways in which internalized norms and prosocial emotions interact to bring about (and sustain) cooperative behavior.

The third conclusion of this chapter is that the abovementioned models are useful to pursue the goal of explanatory progress. Yet the existing behavioral treatments still do not offer a rather detailed account of how exactly emotion and higher order cognition might be related in a causal chain of events that produce manifest choice behavior. In line with the previous two chapters, I suggested that a two-level model of prosocial choice might be useful to fill a blank that restricts attainment of goals of enhanced causal articulation and penetration (section 4). By including insights from psychology and biology into economic analysis, the sketched model aimed to offer a genuinely (causal) explanatory account of cooperative behavior and its puzzles. It is important to stress that the suggested two-level approach aims to be a further development of accounts on the lines of Fehr's and Gintis' re-isolative moves. I do not take a two-level explanatory model of the complex phenomenon of human prosociality as unconditionally superior to any other account of prosocial behavior. However, there are grounds for claiming that the suggested theoretical model can undertake explanatory tasks that most behavioral models cannot. By specifying the chain of mental processes and states (physically realized by particular brain structures and processing activities) by which actual people make judgments and decisions over prosocial (antisocial) acts, a two-level account contributes to causal articulation within the mental realm and causal penetration by digging into the neural realm. In doing this so, a two-level explanatory account provides a rather detailed account of the physical brain entities and activities that enable agents to exercise their capacity to coordinate their mental capabilities for judgments and decisions during social

interaction. [216] Putting somewhat differently, a two-level account is worth developing since it offers an ingenious way of examining whether physical structures with major (causal) roles in decision-making are also involved in robust and significant ways in processing of emotions. In this case, a theoretical model of prosocial behavior that incorporates two levels of analysis might be useful to better track the ways in which emotions, higher order cognition and manifest prosocial behavior are related in an economic world of complex causal dependencies.

This leads me to my final remark. The neurosciences inform social and behavioral scientists about the physical substrate activating and regulating those processes causally relevant to production and maintenance of prosocial behavior patterns. With due concessions, systematic evidence that the brain structures responsible for emotion processing and decision-making are also involved in the human ability to behave in a prosocial (cooperative) manner strengthens the behavioral researcher's plea for exploring further the (causal) significance of emotions to economic analysis of cooperation.

216 According to Antonio Damasio, some brain structures (e.g. amygdala and prefrontal cortex) significant for the agent's capacity to feel social emotions (e.g. anger, shame, guilt, sympathy) also regulate information processing by which human beings can deploy their repertoire of social emotions in order to assess the benefits of prosocial or antisocial action and to quickly select a satisfactory option (2003, p.155).

CHAPTER 8

CONCLUSIONS

> So, it seems right to offer the slightly more courageous prediction that Homo Economicus will become more emotional, by which I mean that economists will devote more attention to the study of emotion.
>
> Richard Thaler, From Homo Economicus to Homo Sapiens, *JEP, 2000*

> The basic need for psychology in economic research consists in the need to discover and analyze the forces behind economic processes, the forces responsible for economic actions, decisions and choices.
>
> George Katona, *Psychological Analysis of Economic Behavior*, 1977

1. Summary

This monograph is an attempt to explore the benefits from interdisciplinary research done by behavioral economists, psychologists and neuroscientists. Its starting point is the vision that, unlike most practicing economists, contemporary behavioral researchers seem to pursue the task of genuine scientific explanation as one that provides knowledge of a phenomenon's causal history (i.e. it identifies the chain of events that produce the *explanandum*) and therefore yields understanding of its place in the causal structure of the world. The current work focused on the investigation of a controversial philosophical issue - the explanation of decision-making and the non-negligible roles emotions may play in it. This chapter briefly explains what I tried to accomplish in the previous chapters and why.

Instead of relying on prepackaged ideas from philosophy of science (and philosophy of economics), I opted for investigating the actual practice of a research field called behavioral economics. The latter is here interpreted as a research agenda that recognizes that the standard economic theory of choice failed to predict and to explain economically

relevant patterns of behavior within three important domains of economic analysis, viz. intertemporal choice, decision under risk and prosocial choice.

This monograph is premised on the idea that behavioral economists engage in amending analytical models/theories of choice so as to improve the quality of predictions and explanations of behavior patterns in the real economic world, where individuals have limited knowledge and constrained computational facilities to deal with uncertainties about the future, risky prospects (and their consequences) including as well as other sources of interpersonal and intrapersonal conflict (chapters 2). With this in mind, I put forward the idea that behavioral economists make claims about the nature and purposes of theorizing and explanation of decision-making behavior that are consistent with a philosophical doctrine called scientific realism (chapter 3). I have suggested that behavioral economists can be interpreted as realists about models/theories of choice provided that they try to come up with improved representations of the complex phenomenon of decision-making by means of manipulations that serve to isolate those major elements in production of the explanandum phenomenon under study from the rest of the world. In so doing, these reformed accounts of choice behavior aim to improve the explanatory and predictive capabilities of economic analysis.

In line with a realist interpretation of behavioral theorizing, theoretical innovations were thought to arise in response to researchers' doubts about whether one or more explanatory factors constituting the conventional choice framework is/are sufficient or even necessary for an account of decision-making behavior that meets its scientific purposes of prediction and/or explanation. This monograph showed that (i) economists and psychologists' detection of choice anomalies indeed motivated the development of reformist strategies, i.e., amended models of choice that employ an incremental strategy or a process description approach, and (ii) such theoretical innovations can indeed be conducive to progress as scope expansion and/or as causal articulation and penetration. In particular, important amended models of choice (e.g. projection heuristic model, affect heuristic model and strong reciprocity model) are regarded as de-isolative or re-isolative moves based on the recognition that the inclusion of a class of explanantia items called

emotions make a significant difference for reducing the gap between predicted and actual behavior within the domains of choice over time, decision under risk and prosocial choice. Furthermore, models of choice describing how emotions guide actual decision-making allow for improved understanding of how actual behavior (anomalies included) comes about. This suggests that it's about time that economic analysis takes the causal roles of emotions seriously.

The main claim of this work is that emotions constitute non-negligible mental events in the production of actual decision-making and therefore ought to be incorporated explicitly in a reformed economic approach to behavior, despite some analytical tractability challenges that they may pose to us. This monograph offered empirical evidence suggesting that models of intertemporal choice, decision under risk and prosocial choice that incorporate affective processes are better in identifying non-negligible driving forces behind actually observed choice behavior than those accounts that fail to do so.

Even though I find the available theoretical innovations through de-isolation and re-isolation very fruitful and conducive to explanatory progress in behavioral economic analysis, there are still some blanks to be filled. In the present work, I focused attention on a particular blank; the problem of specifying how emotions participate in a causal chain of mental events that produce economically important patterns of choice behavior. My point of departure is the economic methodologist's vision that economists ought to address explicitly the difficult issue of causation and to work hard to come up with an explanation that uncovers the causal trajectory through which the phenomenon under study happens (Hausman 2001). Inspired by Simon's theoretical model of choice, I tentatively suggested an explanatory account that purports to deal with the complex task of (causal) explanation, seeking to enhance explanatory progress as increased causal articulation. The novelty of this theoretical account is that it offers a description of the decision machinery with two levels of analysis, mental and neural (chapter 4). What motivated me to descend one level of analysis down is the vision that a physical brain description of the complex machinery for human thinking and decision-making informs

us about those mental processes that are causally significant for economically relevant patterns of behavior, including some apparent choice anomalies within the domains of intertemporal choice, decision under risk and prosocial choice.

I argued that an account of decision-making behavior that includes a brain level of analysis may add to the explanatory power of economic analysis, since it offers an additional way of examining whether (and how) certain mental properties and events (i.e. explaining items) play significant and robust roles in the (causal) production of the explanandum under study. It is important to stress that the proposed theoretical model of decision-making behavior is committed to the vision that mental properties, states and phenomena supervene on physical brain properties and states. It suggests that including a neural level of analysis into an explanatory account of decision-making is worthwhile in that it allows for genuinely scientific understanding of how economically relevant mental phenomena (e.g. emotion and higher order cognition) and physical events (e.g. manifest choice behavior) are related in a complex world of causal dependencies.

The two-level model of behavior allows for an improved understanding of why emotions and feelings are non-negligible components of the machinery for human thinking and decision-making in the real economic world. By describing those brain structures and processing activities that execute physically those mental processes causally relevant to manifest choice behavior, a two level account informs us that emotions amount to specialized systems that enable individuals to exercise crucial inferential capabilities to decision-making like automatic detection of a choice task calling for immediate behavioral response, selective search for choice alternatives that meet their perceived goal aspirations and affect-driven selection of a satisfactory course of action.

In order to show how my schematic explanatory account of decision-making might add to the explanation of puzzling instances of behavior, I applied it to three domains of economic analysis: intertemporal choice, decision under uncertainty and risk and prosocial choice. Based on a study of sequences of models of choice, chapters 5, 6 and 7 confirmed the hypothesis that behavioral models with more explicit psychological

assumptions (than the standard choice framework) are de-isolative or re-isolative moves towards improved predictive and explanatory improvements in economic analysis. Such theoretical innovations have emerged in response to economists' doubts about whether the isolated explaining items of the conventional models of choice are sufficient and even necessary to the development of an account that resolves empirical regularities (taken as anomalies) and offers knowledge of the causal history of important patterns of choice behavior. Additionally, I showed that a two-level account of decision-making behavior can be applied to various domains of human affairs, and that it adds to the explanation of how actual choice behavior (and its puzzles) may arise in the real world.

One important finding of the present work is that a two-level account of choice behavior is unifying in that it invokes the same information processing activities performed by specialized brain structures to explain instances of choice behavior in three diverse domains, that is, it offers a description of those emotional processes (and higher order cognitive processes) with major roles in the causal production of three distinct (and often puzzling) kinds of choice behavior, viz. choice over time, decision under risk and prosocial (cooperative) acts. If this is so, it might also contribute to the pursuit of explanatory progress as unification.

A two-level description of how the decision machinery operates to bring about behavior is worthwhile since it throws extra light on the potential explanatory gains associated with theorizing strategies of re-isolation. I have tried to demonstrate that a two-level account contributes to the development of empirically grounded models (e.g. projection bias model, affect heuristic model and model of prosocial choice with strong reciprocity and emotion driven-altruistic punishment) that aim to understand how emotion, cognition and actual choice behavior are related in a complex economic world of causal (inter)dependencies.

2. Afterthoughts and agenda for future research

The current work offered a metatheoretical analysis of the explanation of decision-making behavior that indicates that 'there is more realism in economics than easily meets

the eye' (Mäki 2002, p. 91). Unlike the standard view that economists endorse the metatheoretical doctrine called theoretical instrumentalism, this monograph shows that some behavioral economists make claims about theorizing and explanation that are consistent with a realist (anti-instrumentalist) standpoint. Important figures like Matthew Rabin, George Loewenstein and Colin Camerer seem to be (moderate) realists about theorizing and explanation since they seek to develop models of choice that might truly represent the way the economic world works.

Even though this monograph focused on discussion and development of a satisfactory explanation of decision-making behavior and the role emotions may play in that, it is obvious that new studies are needed to clarify the terms of the debate about the issues of explanation and causation. This is an important challenge posed to those interested in better understanding the prospects of explanatory progress as scope expansion and causal articulation and its link with predictive improvements in economics.

The present monograph advanced the argument that a two-level account of choice behavior offers a strategy to track mental and neural properties and events with major roles in explaining decision-making behavior and the central role emotions may play in it. The current work puts forth an explanatory account of decision-making behavior that is committed to the view that specialized mental properties, events and states supervene on brain properties, neural structures and processing activities causally sufficient to the occurrence of choice behavior. Yet further research is required so as to come up with an account of behavior that offers a detailed description of the causal and other dependencies between the levels of mind and brain. Finally, I leave for future research the effort of giving a formally tractable treatment to an account of decision-making behavior at the levels of mind and brain and its implications.

BIBLIOGRAPHY

Adolphs, Cahill, Schul and Babinsky (1997) "Impaired Memory for Emotional Stimuli Following Bilateral damage to the Human Amygdala," *Learning and Memory*, 4, pp.291-300

Adolphs, R. (2003) "Cognitive neuroscience of human social behavior," *Nature Reviews/ Neuroscience*, vol. 4, March, pp. 165-178

Adolphs, R. and A. Damasio (2001) "The Interaction of affect and cognition: a neurobiological perspective," In J. Forgas (ed.), *Handbook of affect and social cognition,* Hillsdale: Erlbaum.

Adolphs, R., D. Tranel and A. Damasio (1994) "Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala," *Nature*, 372, pp. 669-672

Adolphs, R., D. Tranel and A. Damasio (1998) "The human amygdala in social judgment," *Nature*, 393, pp. 470-474

Adolphs, R., H. Damasio, D. Tranel, G. Cooper and A. Damasio (2000) "A role for somatosensory cortices in the visual recognition of emotion as revealed by 3-D lesion mapping. *Journal of Neuroscience*, 20, pp. 2683-2690

Ainslie, G. and R. J. Hernnstein (1981) "Preference Reversal and Delayed Reinforcement," *Animal Learning and Behavior*, 9, 4, pp. 476-482

Ainslie, G. and V. Haendel (1983) "The Motives of the Will," In: E. Gotthei, K. Druley. T. Skolda and H. Waxman (eds.) *Etiology aspects of alcohol and drug abuse, Springfield*, Il. Charles Thomas

Akerlof, G and R. Dickens (1982) "The Economic Consequences of Cognitive Dissonance," *American Economic Review*, vol. 72, 3, pp. 307-319

Alkahami and Slovic (1994) "A psychological study of the inverse relationship between perceived risk and perceived benefit," *Risk Analysis*, 14, 6, pp. 1085-1096

Allais, M. (1953) "The foundations of a Positive Theory of Choice Involving Risk and Criticism of the Postulates and Axioms of the American School," In: Allais and Hagen, *Expected Utility Hypotheses and the Allais Paradox*. D. Reidel

Anderson, S., A. Bechara, D. Tranel and A. Damasio (1999) "Impairment of social and moral behavior related to early damage in human prefrontal cortex," *Nature Neuroscience*, 2, pp.1032 -1037

Andreoni, J. (1988) "Privately provided public goods in a large economy: the limits of altruism". *Journal of Public Economics*, 35, February, pp. 57-73

Andreoni, J. (1989) "Giving with Impure Altruism: applications to charity and Ricardian equivalence". *Journal of Political Economy*, 97, pp. 1447-1458

Andreoni, J. (1990) "Impure altruism and donations to public goods: a theory of warm-glow giving". *The Economic Journal*, 100, June, pp. 464-477

Angeletos, G., D. Laibson, A. Repetto, J. Tobacman and S. Weinberg (2001) "The Hyperbolic Discounting Consumption Model: calibration, simulation and empirical evaluation". *Journal of Economic Perspectives*, 15, 3, pp-47-68

Archibald, G. H. Simon and P. Samuelson (1963) "Problems of Methodology: Discussion". *American Economic Review Papers and Proceedings*, vol. 53, no. 2, May, pp. 227-236

Aristotle (1941) *The Complete Works of Aristotle: Rhetoric.* Edited by R. McKeon, New York: Random House

Atran, S. (2001) "A Cheater-Detection Module? Dubious Interpretations of the Wason Selection Task", *Evolution and Cognition*, 7, 2, pp. 187-192

Axelrod, R. (1984) *The Evolution of Cooperation*, New York: Basic Books

Barkow, L. Cosmides and J. Tooby (1992) The Adapted Mind: evolutionary psychology and the generation of culture. New York: Oxford University Press

Batson C. D., K. O'Quin, J. Fultz, M. Vanderplis, and A. Isen (1983) "Influence of self-reported distress and empathy on egoistic versus altruistic motivation to help," *Journal of Personality and Social Psychology*, 45, pp. 706-718

Batson, C. D. (1982) "Prosocial motivation: is it ever truly altruistic?" *Advances in Experimental Psychology*, 1, pp. 199-218

Batson, C. D. and J. Coke (1981) "Empathy: a source of altruistic motivation for helping". In: J. Rushton and R. Sorrentino (eds.) *Altruism and Helping Behavior,* Hillsdale: Lawrence Erlbaum

Batson, C. D., B. Duncan, P. Ackerman, P. Buckey and K. Birch (1981) “Is empathetic emotion a source of altruistic motivation?” *Journal of Personality and Social Psychology*, 40, pp. 290-302
Baumol, W. (1951) “The Neuman-Morgenstern Utility Index: an ordinalist view,” *Journal of Political Economy*, vol. 59, 1, pp.61-66
Bechara, A., A. Damasio, H. Damasio, S. Anderson (1994) “Insensitivity to future consequences following damage to prefrontal cortex”. *Cognition*, 50, pp. 7-15
Bechara, A., D. Tranel and H. Damasio (2000) “Characterization of the decision-making deficits of patients with ventromedial prefrontal cortex lesions,” *Brain*, 123, pp. 2189-2202
Bechara, A., D. Tranel, H. Damasio and A. Damasio (1996) “Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex,” *Cerebral cortex*, 6, pp. 215-225
Bechara, A., H. Damasio, A. Damasio (2000) “Emotion, decision-making and the orbitofrontal cortex,” *Cerebral Cortex*,10, pp. 295-307
Bechara, A., H. Damasio, A. Damasio, and G. Lee (1999) ”Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making,” Journal of Neuroscience, 19, pp. 5473-5481
Bechara, A., H. Damasio, D. Tranel, A. Damasio (1997) “Deciding Advantageously Before Knowing the Advantageous Strategy”, *Science*, 275, pp. 12293-1294
Bechtel, W. and R. Richardson (1993) *Discovering Complexity: decomposition and localization as strategies in scientific research,* Princeton: Princeton University Press
Becker, G. (1976) *The Economic Approach to Human Behavior*. Chicago: University of Chicago Press
Becker, G. (1996) *Accounting for Tastes*. Chicago: University of Chicago Press
Bell, D. (1982) "Regret Theory in Decision-Making under Uncertainty". *Operations Research*, 30, 5, pp.961-981
Bell, D. (1985) "Disappointment in Decision-Making under Uncertainty". *Operations Research*, 33, 1, pp.1-27
Bentham, J. (1781/1789) Na Introduction to the Principles of Morals and Legislation [Online] available from http://www.la.utexas.edu/research/poltheory/bentham/ipml/ipml.toc.html; accessed on July 10th 2006; Internet
Benzion, U., A. Rapoport and J. Yagil (1989) "Discount Rates Inferred from Decisions: na experimental study. *Management Science*, 35, pp. 270-284
Bernoulli, D. (1738/1954) "Exposition of a New Theory of Measurement of Risk" *Econometrica*, vol.22, no.1, Jan., pp.23-36
Berthoz, S., J. Armony, R. Blair and R. Dolan (2002) "A fMRI study of intentional and unintentional (embarassing) violations of social norms," *Brain*, 125, pp. 1696-1708
Binmore, K , A Shaked and J. Sutton (1985) "Testing Non-Cooperative Bargaining Theory: a preliminary study", *American Economic Review*, vol. 75, no. 5, Dec., pp. 1178-1180
Binmore, K. and L. Samuelson (1995) "Learning to be Imperfect: the ultimatum game," *Games and Economic Behavior*, pp. 856-890
Black, R., A. W. Coats and C. Goodwin, eds. (1973) *Marginal Revolution in Economics: interpretation and evaluation.* Durham: Duke University Press
Blaug, M. (1972) “Was There a Marginal Revolution?” *History of Political Economy*, fall, 4, pp.269-280
Blaug, M. (1997) *Economic Theory in Retrospect,* Cambridge: Cambridge University Press, 5th edition
Blaug, M. (2001) “No History of Ideas, Please, We’re Economists”. The Journal of Economic Perspectives, vol. 15, no.1, winter, pp. 145-164
Böhm-Bawerk, E. (1889/1923) *The Positive Theory of Capita,* New York: Stechert
Bosman, R. and F. van Winden (1999) “Emotional hazard in a Power to Take Game”, *The Economic Journal*, 112, pp. 147-169
Bouwel, J. van (2003) ‘Towards a Framework for Pluralism in Economics”, Post Autistic Economics Review, Issue no. 31, May, 2005, article 3. [Online] available from **http://www.paecon.net/PAEReview/pluralism/VanBouwel30.htm**; accessed on July 10th 2006; Internet
Bouwel, J. van (2004) “Explanatory pluralism in economics: against the mainstream? Philosophical Explorations, 7, 3, pp. 299-315
Bowles S., E. Fehr and H. Gintis (2003) "Strong Reciprocity May Evolve With or Without Group Selection". *Theoretical Primatology*, vol.1, Dec., issue 12, supplement
Bowles, S. and H. Gintis (2002) “The Economics of Shame and Punishment,” Santa Fe Institute Working Paper, Available at http://www-unix.oit.umass.edu/~gintis

Bowles, S. and H. Gintis (2003a) "Prosocial Emotions," Santa Fe Institute Working paper # 02-07-028 [Online] available from http://www.umass.edu/preferen/gintis/prosoc.pdf; accessed on July, 10, 2006; Internet
Bowles, S. and H. Gintis (2003b) "The Origins of Human Cooperation," In P. Hammerstein (ed.) *Genetic and Cultural Evolution of Cooperation*, MIT Press
Browne and Hoyt (2000) – about experiments on decision-making under risk
Buss, D. (1999) *Evolutionary Psychology: the new science of the mind*. Allyn & Bacon
Caldwell. B. (1986) "Economic Methodology and behavioral economics: an interpretive history", in *Handbook of Behavioral Economics*, B. Gilad ans S. Kaish (eds.) Greenwich: JAI Press
Camerer, C. (1995) "Individual Decision-Making" In J. Kagel and A. Roth (eds.), *The Handbook of Experimental Economics*. Princeton: Princeton University Press
Camerer, C. (1998) "Bounded Rationality in Individual Decision-Making", California Institute of Technology, May, Social Science Working Paper 1029
Camerer, C. and G. Loewenstein (2004) "Behavioral Economics: past, present and future". In *Advances of Behavioral Economics,* C. Camerer, G. Loewenstein and M. Rabin (eds.), Princeton: Princeton University Press.
Camerer, C., G. Loewenstein and D. Prelec (2004) "Neuroeconomics: why economics needs brains," Scandinavian Journal of Economics, vol. 16, no.3, Sept, pp. 555-579
Camerer, C., G. Loewenstein and D. Prelec (2005) "Neuroeconomics: how neuroscience can inform economics", *Journal of Economic Literature*, vol. XLIII, March, pp. 9-64
Cardinal, R., D. Pennicott, C. Sugathapala, T. Robbins and B. Everett (2001) "Impulsive choice induced in rats by lesions of nucleus accumbens", *Science*, 292, pp. 2499-2501
Carrillo, J. and T. Mariotti (2000) "Strategic Ignorance as a Self-Disciplining Device," *Review of Economic Studies*, 67, 3, pp. 529-544
Carruthers, P. (2003) *The Nature of the Mind: an introduction*. London: Routledge
Casscells, W., H. Schoenberger and T. Grayboys (1978) "Interpretations by physicians of laboratory results", *New England Journal of Medicine*, 299, pp. 999-1000
Christensen-Szalanski, J.J.J (1984) "Discount Functions and the Measure of Patient's Values: women's decisions during childbirth" *Medical Decision-Making*, 4, pp.47-58
Churchland, P. (1981) "Eliminative Materialism and the Propositional Attitudes" *Journal of Philosophy*, LXXVIII, pp.67-90
Churchland, P. (1986) *Neurophilosophy: toward a unified mind-brain*. Cambridge: MIT Press
Churchland, P. (1996) "Feeling Reasons," In: P. Churchland and P. Churchland, *On the Contrary*. Cambridge: MIT Press
Churchland, P. and T, Sejnowski. (1992) *The Computational Brain,* Cambridge: MIT Press
Coats, A. (1976) "Economics and Psychology: the death and resurrection of a research programme," In S. Latsis (ed.) *Method and Appraisal in Economics*, Cambridge: Cambridge University Press
Colinsk, J. (1996) "Why Bounded Rationality?" *Journal of Economic Literature*, 34, pp.669-700
Collard, D. (1978) *Altruism and economy: a study in non-selfish economics*. New York: Oxford University Press
Cosmides, L. and J. Tooby (1992) "Cognitive Adaptations for Social Exchange". In J. Barkow, L. Cosmides and J. Tooby (eds.) *The Adapted Mind*, New York: Oxford University Press
Cosmides, L. and J. Tooby (1996) "Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty" *Cognition*, 58, pp.1-73
Cosmides, L. (1989) "The logic of social exchange: has natural selection shaped how humans reason? Studies with Wason selection task," *Behavioral and Brain Sciences*, 12, pp. 36-37
Cosmides, L. and J. Tooby (2000) "Evolutionary Psychology and the Emotions," In: M. Lewis and J. Haviland (eds.) *Handbook of the Emotions*, 2nd edition, pp. 91-115
Cowan, R. and M. Rizzo (1996) "The Genetic-Causal Tradition and Modern Economic Theory", Working paper. Available http://www.cgl.uwaterloo.ca/~racowan/cause.html, accessed July 11th, 2005
Damasio, A. (2003) *Looking for Spinoza: joy, sorrow and the feeling brain.* New York: Harcourt, Inc.
Damasio, A. (1994) *Descartes' Error,* New York: Putnan
Damasio, A. (1999) *The feeling of what happens: Body and emotion in the making of consciousness,* New York: Harcourt.

Darley and Latané (1968) “Bystander intervention in emergencies: diffusion of responsibility.” *Journal of Personality and Social Psychology*, 8, pp. 377-383
Davenport, H. (1896/1968) *Outlines of Economic Theory,* New York: A. M. Kelley
Davidson, D. (1970) “Mental Events,” In: L. Foster and J. Swanson (eds.) *Experience and Theory*, University of Massachusetts Press
Davidson, R., K. Scherer and H. Goldsmith (2003) *Handbook of the Affective Sciences,* New York: Oxford University Press
Dawes, R., J. van der Krag and J. Orbell (1986/1990) “Cooperation for the benefit of us – not me, or my conscience”, In: J. Mansbridge (ed.) *Beyond Self-Interest*. Chicago: University of Chicago Press
Dickinson, C. (1919) “The Relations of Recent Psychological Developments to Economic Theory” *Quarterly Journal of Economics*, vol. 33, no. 3, May, pp. 377-421
Dow, S. (1997) “Methodological Pluralism and Plurality of Methods”. In A. Salant and E. Screpanti (eds), *Pluralism in Economics: new perspectives in history and methodology*. Cheltenham, UK. Edward Elgar.
Downey, H. (1910) “The Futility of Marginal Utility,” *Journal of Political Economy*, April, 18, 4, pp. 253-268
Earl, P. (1980) “Economics and Psychology: a survey”. *The Economic Journal*, vol. 100, no. 402, Sep., pp.718-755
Edgeworth, F. (1879) “The Hedonical Calculus”, *Mind*, vol. 4, Issue 15, July, pp. 394-408
Edgeworth, F. (1881) *Mathematical Psychics: an essay on the application of mathematics to moral sciences,* London: Kegan Paul
Edwards, W. (1954) “The Theory of Decision-Making”, *Psychological Bulletin*, 41, pp. 380-417
Edwards. W. (1961) “Behavioral Decision Theory,” *Annual Review of Psychology*, 12, pp. 473-498
Eisenberg, N. (1986) *Altruistic Emotion, Cognition and Behavior,* Hillsdale: Lawrence Erlbaum
Ekman P., R. Levenson and W. Friesen (1983) “Autonomic nervous system activity distinguishes among emotions,” *Science*, 221, pp. 1208-1210
Ellsberg, D. (1961) “Risk, Ambiguity and the Savage Axioms”, *Quarterly Journal of Economics* 75, pp. 643-669
Elster, J. (1983) *Explaining Technical Change*. Cambridge: Cambridge University Press
Elster, J. (1985) “The nature and scope of rational choice explanations,” in E. LePore and B. McLaughlin (eds.) *Actions and Events: perspectives on the philosophy of Donald Davidson*. Oxford: Blackwell
Elster, J. (1996) “Emotions and Rationality”, *The Economic Journal*, Sept. 106, 438, pp.1386-1397
Elster, J. (1998) “Emotions and Economic Theory,” *Journal of Economic Literature,* 36, pp. 47-64
Elster, J (1999a) *Alchemies of the Mind: Rationality and the Emotions*. Cambridge, MA: Cambridge University Press.
Elster, J. (1999b) *Strong Feelings: emotion, addiction and human behavior*. Cambridge: MIT Press
Fehr, E. and K. Schmidt (1999) “A Theory of Fairness, Competition and Cooperation” *Quarterly Journal of Economics,* vol. 114, no. 3, pp. 817-868
Fehr, E. and K. Schmidt (2000) “Fairness, Incentives and Contractual Choices”, European Economic Review, vol. 44, no. 2, pp. 393-421
Fehr, E. and S. Gächter (2000) “Cooperation and Punishment in Public Goods Experiments”, *American Economic Review*, 90, pp. 980-994
Fehr, E. and S. Gächter (2002) “Altruistic Punishment in Humans”, Nature, 415, Jan., pp.137-140
Fehr, E., U. Fischbacher and S. Gächter (2000) “Strong reciprocity, human cooperation and the enforcement of social norms,” *Human Nature*, 13, pp. 1-25
Fessler, D. and Haley (2001/ 2003) “The Strategy of Affect: emotions in human cooperation”. In: P. Hammerstein (ed.) *Genetic and Cultural Evolution of Cooperation*. Cambridge: MIT Press.
Festinger, L. (1957) *A Theory of Cognitive Dissonance,* Palo Alto: Stanford University Press
Fiedler, K. (1988) “The Dependence of the Conjunction Fallacy on Subtle Linguistic Factors”, *Psychological Research*, 50, pp.123-129
Fine, C., J. Lumsden and R. Blair (2001) “Dissociation between theory of mind and executive functions in a patient with early left amygdala damage”, Brain, 124, pp. 287-298
Finucane, M., A. Alkahami, P. Slovic and S. Johnson (2000) “The Affect Heuristic in Judgments of Risks and Benefits”, *Journal of Behavioral Decision-Making*, 13, pp. 1-17
Fischbacher, U., S. Gächter and E. Fehr (2001) “Are people conditionally cooperative? Evidence from a public goods experiment,” *Economics Letters*, 71, pp. 397-404

Fischhoff, B., P. Slovic, S. Lichtenstein, S. Reed and B. Combs (1978) "How safe is safe enough? A psychometric study towards technical risks and benefits", *Policy Sciences*, 9, pp. 127-152
Fishburn, P. (1970) *Utility Theory for Decision-Making*. New York: Wiley
Fisher, I. (1892/1965) *Mathematical Investigations in the Theory of Value and Price*, New York: A. M. Kelley
Fletcher, P., C. Frith, P. Grasby, T. Shallice, R. Frackowiak and R. Dolan (1995) "Brain Systems for Encoding and Retrieval of Auditory-Verbal Memory: An In Vivo Study in Humans". *Brain,* 118, pp.401-416
Fodor, J. (1975) *The Language of Thought*. Cambridge: Harvard University Press
Fodor, J. (1983) *The Modularity of Mind*, Cambridge: MIT Press
Foss, K. and N. Foss (2000) "Theoretical Isolation in Contract Theory: suppressing margins and entrepreneurship," *Journal of Economic Methodology*, 7, pp. 313-339
Frank, R. (1988) *Passions within Reason,* New York: W.W. Norton & Co.
Frank, R. (1992) "Moral Sentiments in Intertemporal Choice," In J. Elster and G. Loewenstein: *Choice over Time*. New York: Russel Sage Foundation.
Frederick, S., G. Loewenstein and T. O'Donoghue (2003/2004) "Time Discounting and Time Preference: a critical review. In: D. Read and R. Baumeister (eds.) *Time and Decision: economic and psychological perspectives on intertemporal choice*. New York: Russel Sage. Also published in C. Camerer, G. Loewenstein and Matthew Rabin (eds.), *Advances in Behavioral Economics,* New York: Russel Sage Foundation, 2004
Frey and Weck-Hanneman (1984) "The Hidden Economy as an Unobserved Variable" *European Economic Review*, 26, pp. 33-53
Frey, B. (1992) *Economics as a science of human behavior: toward a new social science paradigm*, Dordrecht: Kluwer Publishers
Friedman, D. (1998) "Monty's Hall Three Doors: construction and deconstruction of a choice anomaly" *American Economic Review,* vol. 88, no. 4, Sept., pp. 933-46
Friedman, M. (1953) "The Methodology of Positive Economics," In: *Essays of Positive Economics*. University of Chicago Press
Friedman, M. and L. Savage (1948) "The Utility Analysis of Choice Involving Risk," *The Journal of Political Economy*, vol. LVI, no. 4, Aug., pp. 279-304
Friedman, M. and L. Savage (1952) "The Expected Utility Hypothesis and the Measurability of Utility", *Journal of Political Economy*, 60, 6, pp. 463-474
Ganzach, Y. (2001) "Judging Risk and Return of Financial Assets", *Organizational Behavior and Human Decision Processes*, 83, pp. 353-370
Gazzaniga, M., R. Ivry and G. Mangun (1998) *Fundamentals of Cognitive Neuroscience,* New York: W.W. Norton
Gibbard, A. (1992) *Wise Choice, Apt Feelings: a theory of normative judgment*. Oxford: Clarendon Press
Gigerenzer, G. (1991) "How to make cognitive illusions disappear: beyond heuristics and biases," *European Review of Social Psychology*, 2, pp. 83-115
Gigerenzer, G. (1996) "On Narrow Norms and Vague Heuristics: a reply to Kahneman and Tversky (1996)," *Psychological Review*, vol. 103, no. 3, pp. 592-596
Gigerenzer, G. (1998) "Ecological Intelligence: an adaptation for frequencies". In. D. Cummins and C. Allen (eds.) *The Evolution of Mind*, New York: Oxford University Press
Gigerenzer, G. (2001) "The Adaptive Toolbox" "In: G. Gigerenzer and R. Selten (eds.), *Bounded Rationality: the adaptive toolbox*. Cambridge, MA. MIT Press
Gigerenzer, G. and D. Goldstein (1996) "Reasoning the fast and frugal way: models of bounded rationality," *Psychological Review*, 103, 4, pp. 650-669
Gigerenzer, G. and P. Todd (1999) "Fast and frugal heuristics: the adaptive toolbox". In: G. Gigerenzer, P. Todd and the ABC Research Group (eds.) *Simple heuristics that makes us smart,* New York*:* Oxford University Press.
Gigerenzer, G. and R. Selten (2001) *Bounded Rationality: the adaptive toolbox*. Cambridge, MA. MIT Press
Gigerenzer, G., P. Todd and the ABC research group (1999) *Simple heuristics that make us smart*. New York: Oxford University Press

Gilad, B. ans S. Kaish, eds. (1986) *Handbook of Behavioral Economics*, Greenwich: JAI Press

Gilbert, D., E. Pinel, T. Wilson, S. Blumberg and T. Wheatley (1997) "Immune effect: a source of durability bias in affective forecasting" *Journal and Personality and Social Psychology*, 75, pp. 617-638
Gintis, H. (2000) "Strong Reciprocity and Human Sociality", *Journal of Theoretical Biology*, 206, pp. 169-179
Gintis, H. (2003) "Solving the Puzzle of Prosociality", *Rationality and Society*, 15, 2, May, pp. 155-187
Glazer, A. and K. Konrad (1996) "A signaling explanation for charity", American Economic Review, vol.86, no. 4, pp. 1019-1028
Glennan, S. (2002) "Rethinking Mechanistic Explanation". *Philosophy of Science*, 69, pp. S342- S353
Goldstein, W. and R. Hogarth (1997) *Research on Judgment and Decision-Making: currents, connections and controversies*, New York: Cambridge University Press
Graham, G. (2005), "Behaviorism". In E. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, [Online] available from http://plato.standford.edu/entries/behaviorism; accessed on July 10, 2006; Internet
Granger, C. (1969) "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods" *Econometrica*, vol. 37, Issue 3, pp. 424-438
Greenberg, J. (1990) "Employee Theft as a Reaction to Underpayment Inequality: the hidden cost of pay cuts". *Journal of Applied Psychology*, 73, pp.556-68
Greene, J. and J. Haidt (2002) "How (and where) does moral judgment work?" *Trends in Cognitive Sciences*, 6, 12, pp. 517-523
Greene, J., R. Sommerville, L. Nystrom, J. Darley and J. Cohen (2001) "An fMRI investigation of emotional engagement in moral judgment", Science, vol.293, Sept., 14, pp. 2105-2108
Griffiths, P. (1997/1998) *What Emotions Really Are: the problem of psychological categories*. Chicago: University of Chicago Press
Groenewegen, J. and J. Vromen (1996) "A Case for Theoretical Pluralism," in J. Groenewegen (ed.) *Transaction Cost Economics and Beyond* Dordrecht: Kluwer
Guala, F. and L. Mittone (2002) "Experiments in Economics: testing theories vs. the robustness of phenomena" Working paper. [Online] available from http://eprints.biblio.unitn.it/archive/00000328/01/paper02_09.pdf ; accessed on July 10, 2006; Internet
Guala, F. and A. Salanti (2002a) "Model Robustness in old and new growth theory," Working Paper University of Bergamo
Guala, F. and A. Salanti (2002b) "On the Robustness of Economic Models," Working Paper no.8, University of Bergamo
Güth W., R. Schmittberger and B. Schwarze (1982) "An experimental analysis of ultimatum bargaining" *Journal of Economic Behavior and Organization*, 3, pp.367-388
Hamilton, W. (1964) "The Genetical Evolution of Social Behavior I." *Journal of Theoretical Biology* 7, pp. 1-16
Hamilton, W. (1971) "Selection of Selfish and Altruistic Behavior in Some Extreme Models," In J. Eisenberg and W. Dillon (eds.), *Man and Beast: comparative social behavior*. Washington: Smithsonian Institution Press
Hands, W. (2001) *Reflection Without Rules: economic methodology and contemporary science theory*. Cambridge: Cambridge University Press
Hanoch, Y. (2002). "Neither an Angel nor an Ant: Emotion as an aid to bounded rationality". *Journal of Economic Psychology,* 23, 1-25
Harrison, G. (1994) "Expected Utility and the Experimentation". *Empirical Economics*, 19, pp. 223-53
Hausman, D. (1992) *The Inexact and Separate Science of Economics*. Cambridge: Cambridge University Press
Hausman, D. (2001) "Explanation and Diagnosis in Economics," *Revue Internationale de Philosophie*, 55, pp. 311-326
Hausman, D. (2003) "Philosophy of Economics." The Stanford Encyclopedia of Philosophy, edited by N. Zalta. Available http://plato.stanford.edu/archives/win2003/entries/economics/>
Henrich, J. and R. Boyd (2001) "Why people punish defectors: weak conformist transmission can stabilize cstly enforcement of norms in cooperative dilemmas", *Journal of Theoretical Biology*, 208, pp. 79-89
Hertwig, R. and A. Ortmann (2001) "Experimental Practices in Economics: a methodological challenge for psychologists," *Behavioral and Brain Sciences*, 24, 3, pp. 383-403
Hicks, J. and R. Allen (1934) "A Reconsideration of Utility Theory", *Economica* 1, 54, pp. 54-76 and 196-219

Hindriks, F. (2005) "Unobservability, tractability and the battle of assumptions," *Journal of Economic Methodology*, 12, 3, pp. 383-406
Hirschleifer, J. (1987) "On the emotions as guarantors of threats and promises". In. J. Dupré (ed.) *The Latest on the best: essays on evolution and optimality*. Cambridge: MIT Press
Hirschman, A. (1985) "Against parsimony: three easy ways of complicating some categories of economic discourse," *Economics and Philosophy*, 1, pp. 7-21
Hobbes, T. (1651) *The Leviathan: or the matter, form and power of a commonwealth with ecclesiasticall and civil*. *McMaster University Archive of the History /Leviathan.pdf*
Hogarth, R. and M. Reder (1986) *Rational Choice,* Chicago: University of Chicago Press
Holcomb, J. and P. S. Nelson (1989) "An Experimental Investigation of Individual Time Preference," unpublished Working Paper
Hollis, M. and R. Sugden (1993) "Rationality in Action," *Mind*, 102, pp. 1-35
Hoover, K. (1990) "The logic of causal inference: Econometrics and the conditional analysis of causality," *Economics and Philosophy*, 6, 2, pp. 207-234
Hosseini, H. (2003) "The Arrival of Behavioral Economics: from Michigan, or the Carnegie School in the 1950s and the early 1960s", *Journal of Socio Economics*, vol. 32, issue 4, pp. 391-409
Hume, D. (1740) *A Treatise of Human Nature: being an attempt to introduce the experimental method of reasoning into moral subjects*. McMaster University, Archive for the History of Economic Thought, available at: http://socserv2.socsci.mcmaster.ca/~econ/ugcm/3II3/hume/treat.html
Isen, A. and N. Gueva (1987) "The influence of positive affect on acceptable level of risk: the person with a large canoe has a large worry". *Organizational Behavior and Human Decision Processes*, 39, pp.145-54
Isen, A., A. Pratkanis, L. Slovic (1984) "The influence of positive affect on risk preference" Mimeo, 92nd Annual Meeting of the American Psychological Association, Toronto, Canada.
Isen, Nygren and Ashby (1988) "Influence of positive affect on the subjective utility of gains and losses: it is jus not worth the risk". *Journal of Personality and Social Psychology*, 55, pp.710-17
James, W. (1884) "What is an emotion," *Mind*, 9, pp. 188-205. In The *Classics of the History of Psychology*, York University, available at **http://psychclassics.yorku.ca/James/emotion.htm**
James, W. (1890) *Principles of Psychology*, In the Classics of the History of Psychology. York University, available at **http://psychclassics.yorku.ca/James/Principles.htm**
Jevons, S. (1871/1970) *The Theory of Political Economy*. Edited by R. Collison, Harmondsworth, Pelican
Jevons, S. (1881) "Review of Edgeworth's Mathematical Psychics," *Mind*, vol. 6, pp. 581-583
Johnson, E. and A. Tversky (1983) "Affect, generalization, and the perception of risk". *Journal of Personality and Social Psychology*, 45, pp. 20-31
Johnson, E., J. Hershey, J. Meszaros and H. Kunreuther (1993) "Framing, Probability Distortions and Insurance Decisions," *Journal of Risk and Uncertainty*, 7, 1, pp. 35-52. Also reprinted in D. Kahneman and A. Tversky (eds.) *Choices, Values and Frames*, Cambridge University Press, 2000
Kagel, J. and A. Roth (1995) *The Handbook of Experimental Economics*. Princeton: Princeton University Press
Kahneman, D. and A. Tversky (1979) "Prospect Theory: an analysis of decision under risk" *Econometrica*, 47, pp. 263-291
Kahneman, D. and A. Tversky (1986) "Rational Choice and the Framing of Decisions," *Journal of Business*, October, part 2, 59, pp. S251-S278
Kahneman, D. and A. Tversky (1996) "On the reality of cognitive illusions," *Psychological Review*, 103, pp. 582-91
Kahneman, D. and J. Snell (1990) "Predicting Utility," In: R. Hogarth (ed.) *Insights in Decision-Making*, Chicago: University of Chicago Press
Kahneman, D., A. Tversky, J. Knetsch and R. Thaler (1986) "Fairness and the Assumptions of Economics," *Journal of Business, vol.59, no.4, part 2, pp. S285-S300*
Kahneman, D., P. Slovic and A. Tversky (1982) *Judgment under uncertainty: heuristics and biases*. Cambridge: Cambridge University Press
Katona, G. (1951) *Psychological Analysis of Economic Behavior*, New York: McGraw Hill
Katona, G. (1975) *Psychological economics*, New York: Elsevier
Katona, G. (1980) *Essays on Behavioral Economics*, University of Michigan Ann Harbor
Keasey, K. (1984) "Regret Theory and Information: a note". *The Economic Journal*, vol. 94, no. 375, Sept., pp. 645-648

Keltner, D. and J. Gross (1998) "Functional Accounts of Emotions," *Cognition and Emotion*, 13, pp. 467-480

Ketelaar, T. and G. Clore (1997) "Emotions and Reason: the proximate effects and ultimate functions of emotions". In G. Matthews (ed.) *Personality, Emotion and Cognitive Science,* Amsterdam: Elsevier

Ketelaar, T. and P. Todd (2000) "Framing our thoughts: ecological rationality as evolutionary psychology's answer to the frame problem". In H. R. Holcomb (Ed.), *Conceptual challenges in Evolutionary Psychology* III*: Innovative Research Strategies* (pp. 179-211). Norwell, MA: Kluwer.

Kim, J. (1995) *Supervenience and Mind: selected philosophical essays*, New York: Cambridge University Press

Kim, J. (1998) *Philosophy of Mind*, Boulder: Westview Press

Kim, J. (2000) *The Mind in a Physical World: an essay on the mind-body problem and mental causation*, Cambridge: MIT Press

Knight, F. (1921) *Risk, Uncertainty and Profit*. Boston, Ma: Hart, Schaffner & Marx; Houghton Mifflin Company. [Online] available from http://econlib.org/library/Knight/knRUP1.html ; accessed 21 April 2005; Internet

Knight, F. (1925) "Economic Psychology and the Value Problem", *Quartely Journal of Economics*, vol. 39, no. 3, May, pp. 372-409

Koopmans, T. (1949) "Identification Problem in Model Construction," *Econometrica*, vol. 17, no. 2, pp. 125-144

Koopmans, T. (1960) "Stationary ordinal utility and impatience," *Econometrica*, 28, pp. 387-428

Krebs, D. (1982) "Altruism: a rational approach". In N. Eisenberg (ed.), *The Development of Prosocial Behavior*, New York: Academic Press

Kreps, D. (1988) *Notes on the Theory of Choice*. Boulder: Westview Press

Kuilen, G. van and P. Wakker (2004) "Learning in the Allais Paradox" University of Amsterdam Working Paper. [Online] available from http://www1.fee.uva.nl/creed/pdffiles/allais.pdf ; accessed on July 10, 2006; Internet

Kuhn. T. (1962) *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press

LaBar, K. and J. LeDoux (2003) "Emotional learning circuits in animals and humans," In R. Davidson, K. Scherer and H. Goldsmith (eds.), *Handbook of Affective Sciences*. New York: Oxford University Press

Laibson, D. (1994) "Essays in Hyperbolic Discounting". *PhD Dissertation*, MIT

Laibson, D. (2001) "A Cue-Theory of Consumption," *Quarterly Journal of Economics*, 116, pp. 81-119

Laibson, D., A. Repetto, J. Tobacman (1998) "Self-Control and Saving for Retirement," *Brookings Papers on Economic Activity* 1, pp. 91-196

Lakatos, I. and A. Musgrave (1970) *Criticism and the Growth of Knowledge*, Cambridge: Cambridge University Press

Landa, J. and X. Wang (2001) "Bounded Rationality of Economic Man: decision-making under ecological, social and institutional constraints" *Journal of Bioeconomics*, 3, pp. 217-235

Landes, W. and R. Posner (1978) "Altruism in Law and Economics," *American Economic Review Papers and Proceedings*, 68, May, pp. 417-421

Lane, R. and L. Nadel (2002) *Cognitive Neuroscience of Emotion,* New York: Oxford University Press

Lea, S., R. Tarpy and P. Webley, *The Individual in the Economy: a survey of economic psychology,* Cambridge: Cambridge University Press

LeDoux , J (1996/1998) *The Emotional Brain: the mysterious underpinnings of emotional life*. New York: Phoenix

LeDoux, J. (2002) The Synaptic Self: how our brains become who we are. New York: Viking Penguin

Leibenstein, H. (1976) *Beyond Economic Man: a new foundation for microeconomics*. Cambridge: Harvard University Press

Lerner, J. and D. Keltner (2001) "Fear, Anger and Risk." *Journal of Personality and Social Psychology*, 81, pp. 146-159

Lewin, S. (1996) "Economics and Psychology: Lessons For Our Own Day, The Early Twentieth Century", *Journal of Economic Literature*, Sept., vol.34, pp. 1293-1323

Lewis, D. (1983) "New Work for a Theory of Universals," *Australian Journal of Philosophy*, 61, pp. 343-377

Lewis, D. (1986) *On the Plurality of the Worlds*, Oxford; Blackwell
Lewis, D. (1994) "Reduction of the Mind," In S. Guttenplan, *A Companion to the Philosophy of Mind,* Oxford: Blackwell
Loewenstein, G. (1987) "Anticipation and the Valuation of Delayed Consumption," *The Economic Journal*, 97, Sept., pp. 666-684
Loewenstein, G. (1988) "Frames of Mind in Intertemporal Choice," *Management Science*, 34, pp. 200-214
Loewenstein, G. and R. Thaler (1989) "Anomalies: intertemporal choice," *Journal of Economic Perspectives*, vol. 3, no. 4, fall, pp. 181-193
Loewenstein, G. and D. Prelec (1991) "Negative Time Preferences," *The American Economic Review, Papers and Proceedings*, vol. 81, no. 2, May, pp. 347-352
Loewenstein, G. (1996) "Out of Control: visceral influences on behavior". *Organizational Behavior and Human Decision Processes,* 65, pp. 272-92
Loewenstein, G. (1999) "Experimental economics from the vantage point of behavioral economics," *The Economic Journal*, 109, F25-F34
Loewenstein, G. (2000) "Emotions in economic theory and economic behavior" *American Economic Review Papers and Proceedings*, vol. 90, no. 2, May, pp. 426-32
Loewenstein, G. T. O'Donoghue and M. Rabin (2000) "Projection Bias in Predicting Future Utility," Economics Working Paper, University of California at Berkeley. Electronic version available at www.ideas.repec.org
Loewenstein, G. and D. Adler (1995) "A Bias in Prediction of Tastes," *Journal of Personality and Social Psychology*, 46, pp. 961-978
Loewenstein, G. and D. Prelec (1992) "Anomalies in Intertemporal Choice: evidence and interpretation" *Quarterly Journal of Economics*, 107, 2, pp. 573-97
Loewenstein, G. and D. Prelec (1993) "Preferences for Sequences of Outcomes," *Psychological Review*, 100, 1, pp. 91-108
Loewenstein, G. and D. Schkade (1999) "Wouldn't it be Nice? Predicting Future Feelings," In D. Kahneman, E. Diener and N. Schwarz (eds.) *Well-Being: the foundations of hedonic psychology*. New York: Russel Sage Foundation
Loomes, G. and R. Sugden (1982) "Regret Theory: an alternative theory of rational choice under uncertainty," *The Economic Journal*, 92, pp. 805-824
Loomes, G. and R. Sugden (1986) "Disappointment and Dynamic Consistency in Choice under Uncertainty," *Review of Economic Studies*, 53, pp. 271-282
Keynes, J. M. (1936/1967) *The General Theory of Employment, Interest and Money,* London: Macmillan
Kitcher, P. (1989) "Explanatory Unification and the Causal Structure of the World," In P. Kitcher and W. Salmon (eds.) *Scientific Explanation, Minnesota Studies in the Philosophy of Science*, vol. 13, Minneapolis: University of Minnesota Press
Loewenstein, G. and E. Angner (2003) "Predicting and Indulging Changing Preferences," In G. Loewenstein, D. Read and R. Baumeister (eds.) *Time and Decision: economic and psychological perspectives on intertemporal choice*. New York: Russel Sage Foundation
Loewenstein, G. and J. Lerner (2003) "The Role of Affect in Decision-Making," In R. Davidson, K. Scherer and H. Goldmith (eds.) *Handbook of Affective Science*, Oxford: Oxford University Press
Loewenstein, G. and T. O'Donoghue (2004) "Animal Spirits: affective and deliberative processes in economic behavior". Working paper [Online] available from http://www.hss.cmu.edu/departments/sds/faculty/Loewenstein/will7_04.pdf ; accessed on July 10, 2006; Internet
Loewenstein, G., E. Weber, C. Hsee and E. Welch (1998/2001) "Risk as Feelings", *Psychological Bulletin, 127*, 267-286.
Loewenstein, G., T. O'Donoghue and M. Rabin (1999) "Projection Bias in Predicting Future Utility", University of Berkeley Working Paper E00-284. [Online] available from http://elsa.berkely.edu/users/rabin/papers.html; accessed on July 10, 2006; Internet. Also published in the *Quarterly Journal of Economics*, 118, 4, Nov., pp. 1209-1248, 2004
Loomes, G. (1988) "Further Evidence of the Impact of Regret and Disappointment In Choice Under Uncertainty". *Economica*, 55, 217, pp. 47-62
Loomes, G. and R. Sugden (1982) "Regret Theory: an alternative theory of rational choice under uncertainty". *The Economic Journal*, vol. 92, no. 368, Dec., pp. 805-824

Loomes, G. and R. Sugden (1986) “Disappointment and Dynamic Consistency in Choice Under Uncertainty’ *The Review of Economic Studies*, vol. 53, no. 2, April, pp. 271-282

Loomes, G. and R. Sugden (1987) "Testing for Regret and Disappointment in Choice under Uncertainty," *Economic Journal 97* (Supplement), pp. 118-129

Loewenstein, G., B. Shiv, A. Bechara, H. Damasio and A. Damasio (2002) “The Dark Side of Emotions in Decision-Making: when a dysfunctional ventromedial prefrontal cortex can lead to more optimal decisions”. University of Iowa Paper

Macfie, A. (1953) “Choice in Psychology and as Economic Assumption”. *The Economic Journal*, vol. 63, no. 250, June, pp. 352-367

Machamer P., L. Darden and C. Craver (2000) “Thinking About Mechanisms”, *Philosophy of Science*, 67, pp. 1-25

Machina, M. (1982) “Expected Utility Without the Independence Axiom,” *Econometrica*, vol. 50, pp. 277-323

Mäki, U. (1988) “How to combine rhetoric with realism in the methodology of economics”, *Economics and Philosophy* 4, pp. 89-109

Mäki, U. (1989) “On the problem of realism in economics”. Ricerche Economiche, 42, pp.176-198 [Reprinted in B. Caldwell (ed.) *The Philosophy and Methodology of Economics,* Cheltenham: Edward Elgar]

Mäki, U. (1992a) “On the Method of Isolation in Economics”. In C. Dilworth (ed.), *Poznam Studies in the Philosophy of the Sciences and the Humanities*, special issue, 26, pp.319-354

Mäki, U. (1992b) “Friedman and realism,” *Research in the History of Economic Thought and Methodology*, 10, pp.171-195

Mäki, U. (1993a) “Isolation, idealization and truth in economics”. In: B. Hamminga and N. De Marchi (eds.) *Poznam Studies in the Philosophy of the Sciences and the Humanities*, special issue, 38, pp.147-168

Mäki, U. (1993b) “The Market as an Isolated Causal Process: a metaphysical ground for realism”. In: B. Caldwell and S. Boehm (eds.), Austrian Economics: tensions and new developments. Dordrecht: Kluwer Publishers

Maki, U. (1994) “Reorienting the Assumptions Issue”. In R. Backhouse (ed.) *New Directions in Economic Methodology,* London: Routledge

Mäki, U. (1996a) “Scientific realism and some peculiarities of economics” *Boston Studies in the Philosophy of Science*, 169, pp.425-445

Mäki, U. (1996b) “Two Portraits of Economics,” *Journal of Economic Methodology*, 3, 1, pp.1-38

Mäki, U. (1997) “The One World and The Many Theories”. In: A. Salanti and E. Screpanti (eds.) *Pluralism in Economics: new perspectives in history and methodology. Cheltenham*, UK. Edward Elgar.

Mäki, U. (1998a) “Realisticness”. In: J. Davis, W. Hands and U. Mäki, *Handbook of Economic Methodology*. Cheltenham, UK: Edward Elgar

Mäki, U. (1998b) “As if” In: J. Davis, W. Hands and U. Mäki, *Handbook of Economic Methodology*. Cheltenham, UK: Edward Elgar

Mäki, U. (1998c) “Realism”. In: J. Davis, W. Hands and U. Mäki (eds.), *Handbook of Economic Methodology*. Cheltenham: Edward Elgar

Mäki, U. (1998d) “Mechanisms, Models and Free Riders. In R. Backhouse, D. Hausman, U. Mäki and A. Salanti (eds.), *Economics and Methodology: crossing the boundaries*. London: Macmillan.

Mäki, U. (1998e) “Aspects of Realism About Economics”. *Theoria*, 13, pp. 301-313

Mäki, U. (2000) “Kinds of Assumptions and their truth: shaking an un-twisted F-twist”. *Kyklos*, 53, pp.303-322

Mäki, U. (2001a) “Models”. In: International Enclyclopedia of the Social and Behavioral Sciences, vol. 15, pp.9931-9937. Amsterdam Elsevier

Mäki, U. (2001b) “Realism and its opponents,” In: *International Enclyclopedia of the Social and the Behavioral Sciences*, vol. 15, pp.9931-9937. Amsterdam Elsevier

Mäki, U. (2001c) “The way the world works (www): towards an ontology of theory choice,” In U. Mäki (ed.) *The Economic World View: studies in the ontology of economics*. Cambridge, MA. Cambridge University Press

Mäki (2001d) “Explanatory Unification: double and doubtful,” *Philosophy of the Social Sciences*, vol.31, pp. 488-506

Mäki, U. (2003a) "Realism and the Nature of Theory: a lesson from J.H. von Thünen for economists and geographers". Paper presented at the workshop on Scientific Realism and Economic Geography. Erasmus University Rotterdam, June 23, 2003.
Mäki, U. (2003b) "'The Methodology of Positive Economics (1953)' does not give us the methodology of positive economics," *Journal of Economic Methodology*, 10, 4, pp.495-505
Mäki, U. (2004) "Theoretical Isolation and Explanatory Progress: transaction cost economics and the dynamics of dispute", *Cambridge Journal of Economics*, 28, pp. 319-346
Mäki, U. (2005) "Models are Experiments and Experiments are Models," Journal of Economic Methodology, 12, 2, June, pp. 303-315
Manuck, S., J. Flory, M. Muldoon and R. Ferrell (2003) "A neurobiology of intertemporal choice?" In G. Loewenstein, D. Read and R. Baumeister (eds.) *Time and Decision*, New York; Russel Sage Foundation
Marchi, N. de (1973) "Mill and Cairness and the Emergence of Marginalism in England," In C. Black, A. Coats and C. Goodwin (eds.)
Marchionatti, R. and E. Gambino (1997) "Pareto and Political Economy as a Science: methodological revolution and analytical advances in economic theory in the 1890s" *The Journal of Political Economy*, vol. 105, no. 106, pp.1322-1345
Margolis, H. (1984) *Selfishness, Altruism and Rationality*, Chicago: University of Chicago Press
Margolis, H., R. Dawes, A. van de Kragt and J. Orbell (1990) "Cooperation for the Benefit of Us, Not Me or My Conscience," In J. Mansbridge (ed.) Beyond Self-Interest. Chicago: University of Chicago Press
Marr, D. (1982) *Vision: a computational investigation into the human representation and processing of visual information*. New York: W. H. Freeman.
Marshall, A. (1885) *The Present Position of Economics*. London: Macmillan
Marshall, A. (1890/1920) *Principles of Economics*, London: Macmillan
Marshall", A. (1881) "Review of F. Y. Edgeworth's *Mathematical Psychics*", In J. Whitaker, editor (1975) *The Early Economic Writings of Alfred Marshall 1867-90*, vol. 2, Sec Iv.2.11, pp. 265-268 London: Macmillan
Mayr, E. (1988) "Cause and Effect in Biology," In E. Mayr (Ed.), *Toward a New Philosophy of Biology* (pp 24-37), Harvard: Harvard University Press.
McCabe, K., D. Houser, L. Ryan and T. Trouard (2001) "A Functional Imaging Study of Cooperation in Two-Person Reciprocal Exchange." *Proceedings of the National Academy of Science*, 98, 20, pp. 11832-11835
McCabe, K., M. Ringdom and V. Smith (2003) "Sustaining Cooperation in Trust Games" [Online] available from http://www.ices-gmu.net/pdf/materials/415.pdf; accessed on July 10, 2006; Internet, Forthcoming in *The Economic Journal*
McDougall, W. (1910/1926) *Introduction to Social Psychology,* Boston: J. Luce and Co.
McElreath, R., T.H.Clutton-Brock, E. Fehr, D.M. Fessler, E.H. Hagen, P. Hammerstein, M. Kosfeld, M. Milinski, J. Tooby, M.I. Wilson (2001/2003) "The Role of Cognition and Emotion in Cooperation," In P. Hammerstein (ed.) *The Genetic and Cultural Evolution of Cooperation*, Dahlem Workshop Series, Boston: MIT Press
Mele, A. and P. Rawling (2004) *The Oxford Handbook of Rationality,* New York: Oxford University Press
Meeks, G. (1991) *Thoughtful economic man,* Cambridge: Cambridge University Press
Mischel, W. (1974) "Processes in Delay Gratification", In L. Berkowitz *Advances in Experimental Social Psychology*, New York: Academic Press
Mischel, W. Y. Shoda and P. Peake (1988) "The Nature of Adolescent Competences Predicted by Preschool Delay of Gratification," *Journal of Personality and Social Psychology*, 54, 4, pp. 687-696
Mischel, W., Y. Shoda and M. Rodriguez (1992) "Delay of Gratification in Children" In G. Loewenstein and J. Elster (eds.) *Choice over time, New York: Russel Sage*
Mitchell, W. (1914) "Human Behavior and Economics: a survey of recent literature", *Quarterly Journal of Economics*, vol. 29, no. 1, Nov, pp.1-47
Mogenson, G., d. Jones and C. Yim (1980) "From motivation to action: functional interface between limbic system and the motor system". *Progress in Neurobiology*, 14, pp. 69-97
Monroe, K., M. Barton and U. Klingeman (1988/1990) "Altruism and the theory of rational action: rescuers of jews in nazi Europe". *Ethics*, vol. 101, Oct. pp. 103-122
Muramatsu, R. (1999) *Sobre a possibilidade do altruísmo: uma incursão pela análise econômica do comportamento não-egoísta*. MPhil thesis, Department of Economics, University of São Paulo, Brazil

Muramatsu (2002) “Explaining Rational Choice Anomalies by Mechanism”, unpublished
Muramatsu, R. and Y. Hanoch (2005) “Emotions as Mechanism for Boundedly Rational Agents: the fast and frugal way”. *Journal of Economic Psychology*, 26, pp. 201-221
Musgrave, A. (1981) “Unreal assumptions in economic theory: the F-twist untwisted. *Kyklos*, 34, pp. 377-387
Nagel, E. (1963) “Assumptions in Economic Theory”, *The American Economic Review, Papers and Proceedings*, vol. 53, no. 2, May, pp. 211-219
Nelson, A. (2001) “Two models of idealization in economics,” In U. Mäki (ed.) *The Economic World View: studies in the ontology of economics*. Cambridge: Cambridge University Press
Nesse, R. (1990) “Evolutionary Explanations of Emotions,” *Human Nature* vol. 1, no. 3, pp. 261-89
Nesse. R. (2001) “The Evolution of Subjective Commitment,” In R. Nesse (ed.) *Evolution and the Capacity for Commitment*. New York: Russel Sage Press.
Neuman, J. von and O. Morgenstern (1944/1947) *Theory of Games and Economic Behavior*, Princeton: Princeton University Press
Nygren, T., A. Isen and J. Dulin (1996) “The influence of positive affect on the decision rule in risk situations: focus on outcome (and especially avoidance of loss) rather than probability”. *Organizational Behavior and Human Decision Processes*, 66, pp. 59-72
O’Donoghue, T. and M. Rabin (1999) “Doing It Now or Later,” *American Economic Review*, 89, 1, pp. 103-124

O’Donoghue, T. G. Loewenstein and M. Rabin (2000) “Projection Bias in Predicting Future Utility”. Department of Economics Working Paper E00-284, published in the *Quarterly Journal of Economics*, 118, 4, Nov. 2003, pp. 1209-1248
Ochsner, K. and D. Schacter (2003) “Rethinking emotional events: a social cognitive neuroscience approach,” In R. Davidson, K. Scherer and H. Goldsmith (eds.) *Handbook of the affective Sciences*. New York: Oxford University Press
Ochsner, K. and L. Feldman Barrett (2001) “A multiprocess perspective on the neuroscience of emotion,” In T. Mayne and G. Bonnano (eds.) *Emotion: current issues and future directions*. New York; Guilford Press
Ortony, A. and T. Turner (1990) “What’s basic about basic emotions?” *Psychological Review*, 97, pp. 315-331
Papineau, D. (1996) *The Philosophical Naturalism,* Oxford: Blackwell.
Pareto, V. (1896/1964) “Cours d’économie politique”, reprinted in G. Busino, (ed.) *Oeuvres Complètes*, Geneva: Librairie Droz
Pareto, V. (1898/1966) “Comment se pose le problème de l’économie pure”, paper presented at the Societé Stella Conference, reprinted in G. Busino, (ed.) *Oeuvres Complètes*, Geneva: Librairie Droz
Pareto, V. (1900/ 1953/1966) “Sul fenomeno economico (lettera a Benedetoo Croce)”, Giornale degli Economisti, 21, Aug., pp. 139-162, English translation “On the economic phenomenon: a reply to Benedetto Croce”. *International Economic Association Papers*, vol. 2, no. 3, pp. 180-196. Reprinted in G. Busino, (ed.) *Oeuvres Complètes*, Geneva: Librairie Droz
Pareto, V. (1900b/1964) “Sunto di alcuni capitoli di un nuovo trattato di economia pura del prof. Pareto” *Giornale degli Economisti*, March, pp.216-35 and June pp.511-545, reprinted in G. Busino, (ed.) *Oeuvres Complètes*, Geneva: Librairie Droz

Pareto, V. (1901/ 1953/1982) “Sul principio economico,” Giornale degli Economisti, 22, Feb, pp. 131-138 English translation “On the economic principle: a reply to Benedetto Croce”. *International Economic Association Papers*, vol. 2, no. 3, pp. 203-207. Reprinted in G. Busino, (ed.) *Oeuvres Complètes*, Geneva: Librairie Droz
Pareto, V. (1906/ 1971) *Manual of Political Economy*, New York: Augustus M. Kelley

Pareto, V. (1911/ 1955) “Mathematical Economics”, *International Association Economic Papers*, vol. 3, no. 5, pp. 58-102
Peters, E. and P. Slovic (1996) “The Role of Affect and Worldviews as Orienting Dispositions in the Perception and Acceptance of Nuclear Power” *Journal of Applied Social Psychology*, 26, 16, pp. 1427-53
Piatelli-Palmarini (1994) *Inevitable Illusions,* New York: John Wiley and Sons.
Pilliavin, J., J. Dovidio, S. Gaertner and R. Clark (1981) *Emergency Intervention*, New York: Academic Press

Pillutla, M. and K. Murnighan (1996) “Unfairness, Anger and Spite: emotional rejections of ultimatum offers”. *Organizational Behavior and Human Decision Processes*, vol. 68, no. 3, Dec., pp. 208-224
Pinker, S. (2002) *The Blank Slate: the modern denial of human nature*, New York: Viking Penguin
Plutchik, R. (1980) “A psychoevolutionary theory of emotions,” In: R. Plutchik and H. Kellerman (eds.) *Emotion: theory, research and experience*, vol. 1, theories of emotion. New York: Academic
Posner, R. (1993) “Ronald Coase and Methodology” *Journal of Economic Perspectives*, 7, 4, pp. 195-210
Quiggin, J. (1994) “Regret Theory with General Choice Sets”, *Journal of Risk and Uncertainty*, 8, 2, pp.153-162
Rabin, M. (1998) “Psychology and Economics,” *Journal of Economic Literature*, XXXVI, pp. 11-46
Rabin, M. (2002) “A Perspective on Economics and Psychology”, *European Economic Review,* vol. 46, issue 4, pp. 657-685
Rahman, S., B. Sahakian, R. Cardinal, R. Rogers, T. Robbins (2001) “Decision-making and neuropsychiatry,” *Trends in Cognitive Sciences*, 5, pp. 271-277
Railton, P. (2001) “Emotion and rationality” Discussion paper of the Summerschool on Emotion and Rationality, Erasmus University of Rotterdam, the Netherlands.

Ramsey (1926/1931) “Truth and Probability”, In Foundations of Mathematics and Other Philosophical Essays, London: Kegan Paul
Read, D. and B. van Leeuwen (1998) “Predicting Hunger: the effects of appetite on delay and choice”. *Organizational Behavior and Human Decision Processes*, 76, pp. 189-205
Rilling, J., D. Gutman, T. Zeh, G. Pagnoni, G. Berns, C. Kilts (2002) “A neural basis for social cooperation,” *Neuron*, 35, pp. 395-405
Robbins, L (1932/1935) *An Essay on the Nature and Significance of Economic Science*. London: Macmillan)
Roseman, I. and C. Smith (2001) “Appraisal Theory: overview, assumptions, varieties and controversies”. In K. Scherer, A. Schorr and T. Johnstone (eds.), *Appraisal Processes in Emotion: theory, methods and research*. New York: Oxford University Press
Rosenberg, A. (1992) *Economics: mathematical politics or the science of diminishing returns*. Chicago: University of Chicago Press
Rosenberg, A. (1995) *Philosophy of Social Science*. Boulder: Westview Press
Rottenstreich, Y. and C. Hsee (1999/2001) "Money, Kisses and Electric Shocks: the affective psychology of risk” *Psychological Science*, 12, pp. 185-190
Rozin, P. (1976) "The Evolution of Intelligence and Access to the Cognitive Unconscious," In J. Prague and ª Epstein (eds.) *Progress in Psychobiology and Physiological Psychology*. New York: Academic Press

Rubinstein, A. (1988) *Modeling Bounded Rationality*. MIT Press
Rubinstein, A. (2005) “Discussion of Behavioral Economics”. To appear in Advances in Economics and Econometrics: theory and applications, The Ninth Congress. [Online] available from **http://arielrubinstein.tau.ac.il/papers/behavioral-economics.pdf** ; accessed on July 10, 2006; Internet
Ruby, P. and J. Decety (2001) "Effect of subjective perspective taking during simulation of action: a PET investigation of agency". *Nature Neuroscience*, 4, 5, pp. 546-550
Salanti, A. and E. Screpanti (1997) *Pluralism in Economics: new perspectives in history and methodology*, Cheltenham: Edward Elgar
Salmon, P. (1998) “Free-Riding as Mechanism,” In R. Backhouse, D. Hausman, U. Mäki and A. Salanti (eds.), *Economics and Methodology: crossing the boundaries*. London: Macmillan.
Salmon, W. (1984) *Scientific Explanation and the Causal Structure of the World*, Princeton: Princeton University Press
Samuels, R., S. Stich (2004) “Rationality and Psychology”, in A. Mele and P. Rawling (eds.) *The Oxford Handbook of Rationality,* New York: Oxford University Press
Samuels, R., S. Stich and M. Bishop (2002) “Ending the Rationality Wars: how to make disputes about human rationality to disappear”. In: *Commonsene, Reasoning and Rationality*, R. Elio (ed.), Vancouver Studies in Cognitive Science 11. New York: Oxford University Press. [Online] available from http://ruccs.rutgers.edu/ArchiveFolder/Research%20Group/Publications/Wars/wars.html; accessed on July 10, 2006; Internet
Samuels, R., S. Stich and L. Faucher (2002) “Reason and Rationality”, In: *Handbook of Epistemology,* I. Niiniluoto, M. Sintonen and J. Wolenski (eds.), Dordrecht: Kluwer. [Online] available from

http://ruccs.rutgers.edu/ArchiveFolder/Research%20Group/Publications/Reason/ReasonRationality.html; accessed on July 10, 2006; Internet
Samuels, R., S. Stich and P. Tremoulet (1999) "Rethinking Rationality: from bleak implications to Darwinian modules". In: *What is Cognitive Science?* LePore and Z. Pylyshyn (eds.), Oxford: Blackwell. [Online] available from http://ruccs.rutgers.edu/ArchiveFolder/Research%20Group/Publications/Rethink/rethink.html; accessed on July 10, 2006; Internet
Samuels, W. (1997) "The Case for Methodological Pluralism", In Salanti and Screpanti (eds.) *Pluralism in Economics*, Cheltenham: Edward Elgar
Samuelson, P. (1937) "A Note on Measurement of Utility" *The Review of Social Studies*, vol. 4, no. 2, Feb., pp. 155-161
Samuelson, P. (1938a) "A Note on the Pure Theory of Consumer's Behaviour". *Economica*, 5, 17, Feb., pp. 61-71

Samuelson, P. (1938b) "The Numerical Representation of Ordered Classifications and the Concept of Utility" *The Review of Social Studies*, vol. 6, no. 1, Oct., pp. 65-70

Samuelson, P. (1947) *Foundations of Economic Analysis*, Cambridge: Harvard University Press
Samuelson, P. (1948) "Consumption Theory in Terms of Revealed Preference", *Economica*, 15, 60, Nov., pp. 243-253
Samuelson, P. (1952) "Probability, Utility and the Independence Axiom," *Econometrica*, vol. 20, no. 4, Oct., pp. 670-678
Samuelson, P. (1963) "Problems of Methodology – Discussion," *American Economic Review Papers and Proceedings*, 53, pp. 231-236
Satz, D. and J. Ferejohn (1994) "Rational Choice and Social Theory," *Journal of Philosophy*, pp. 71-87
Savage, L. (1954) *Foundations of Statistics*, New York: Wiley
Schkade, D. and D. Kahneman (1998) "Does living in California make people happy? A focusing illusion in judgments of life satisfaction" *Psychological Science*, 9, 5, pp. 340-345
Schoemaker, P. (1982) "The Expected Utility Model: its variants, purposes, evidence and limitations" *Journal of Economic Literature*, vol. 20, no. 2, June, pp. 529-563
Schumpeter, J. (1908) "On the Concept of Social Value". *Quarterly Journal of Economics*, vol. 23, pp. 213-232
Sen, A. (1977) "Rational Fools: a critique of the behavioral foundations of economic theory". *Philosophy and Public Affairs*, vol. 6, no. 4, summer, pp. 317-344
Sen, A. (1990) *On Ethics and Economics*, Oxford: Blackwell
Sen, A. (1993) "Internal consistency of choice" *Econometrica*, 61, pp. 495-521
Shizgal, P. (1999) "On the neural computation of utility: implications from studies of brain stimulation reward". In N. Schwartz, D. Kahneman and E. Diener (eds.) *Well-being: the foundations of hedonic psychology*. New York: Russel Sage Foundation
Sieff. E.; R. Dawes and G. Loewenstein (1999) "Anticipated versus actual responses to HIV test results" *American Journal of Psychology*, 112 (2), pp. 297-311
Simon, H. (1953) "Discussion: decision-making and the theory of organization" *Econometrica*, 21, pp. 348
Simon, H. (1955) "A behavioral model of rational choice," *Quarterly Journal of Economics*, 69, pp. 99-118
Simon, H. (1956) "Rational Choice and the Structure of Environments," *Psychological Review*, 63, pp. 129-138
Simon, H. (1959) "Theories of Decision-Making in Economics and Behavioral Science" American Economic Review, vol. 49, no. 3, June, pp. 253-283

Simon, H. (1962) "The Architecture of Complexity" *Proceedings of the American Philosophical Association*, 106, pp. 467-482
Simon, H. (1967) "Motivational and Emotion Controls of Cognition". *Psychological Review*, vol. 74, no. 1, pp. 29-39
Simon, H. (1978) "Rationality as Process and as Product of Thought" *American Economic Review Papers an*d Proceedings, vol. 68, no. 2, May, pp. 1-16
Simon, H. (1979) "Rational Decision-Making in Business Organization" *American Economic Review*, vol. 69, no. 4, Sept., pp. 493-513
Simon, H. (1983) *Reason in Human Affairs* Stanford: Stanford University Press

Simon, H. (1986) "Rationality in psychology and economics" *Journal of Business*, 59, 4, part II, pp. S209-S224
Simon, H. (1987) "Bounded Rationality". In: J. Eatwell et al, *The New Palgrave: a dictionary of Economics*. London: Macmillan
Simon, H. (1990) "Invariants of Human Behavior," *Annual Review of Psychology*, 41, pp. 1-19

Simon, H. (1992a) "Scientific Discovery as Problem Solving" In M. Egidi & R. Marris (Eds.), *Economics, Bounded Rationality and the Cognitive Revolution*. Aldershot: Edward Elgar.
Simon, H. (1992b) "What is an Explanation of Behavior?" *Psychological Science*
Simon, H.(1997) Models of Bounded Rationality: empirically grounded economic reason, vol. III. Cambridge: MIT Press

Skinner, B. (1953) *Science and Human Behavior*, London: Macmillan

Skinner, B. (1987) "Behaviorism, Skinner On", in *Oxford Companion to the Mind*, New York: Oxford University Press

Slovic, P. and E. Weber (2002) "Perception of Risk Posed by Extreme Events," Discussion Paper [Online] available from http://www.Ideo.columbia.edu/.../meetings/roundtable/white_papers/slovic_wp.pdf; accessed on July 10, 2006; Internet
Slovic, P. (1987) "Perception of Risk" *Science*, 236, pp. 280-285

Slovic, P. and A. Tversky (1974) "Who accepts Savage's axiom?" *Behavioral Science*, 19, pp. 368-373
Slovic, P. B. Fischhoff and S. Lichtenstein (1977) "Behavioral Decision Theory," *Annual Review of Psychology*, vol. 28, pp. 1-39
Slovic, P. B. Fischhoff and S. Lichtenstein (1980) "Facts and fears: understanding perceived risk. In: D. Kahneman, P. Slovic and A. Tversky (eds.) *Judgment under Uncertainty: heuristics and biases*. New York: Cambridge University Press, 1982
Slovic, P., B. Fischhoff and S. Lichtenstein (1984) "Behavioral decision theory perspectives on risk and safety", *Acta Psychologica*, 56, pp. 183-203
Slovic, P., M. Finucane, E. Peters and D. MacGregor (2002a) "The Affect heuristic", in T. Gilovich, D. Griffin and D.Kahneman (eds.), *Intuitive Judgment*. New York: Cambridge University Press
Slovic, p. M. Finucane, E. Peters and D. MacGregor (2002b) "Rational Actors of Rational Fools? Implications of the affect heuristics for behavioral economics", *Journal of Socio Economics*, 31, issue 4, pp. 329-342. [Online] available from http://heuristics.behaviouralfinance.net/affect/Slov02.pdf; accessed on 22 April 2005; Internet
Slovic, P., M. Finucane, E. Peters and D. MacGregor (2003) "Risk as analysis and risk as feelings: some thoughts about affect, reason, risk and rationality". Unpublished paper, [Online] available from http://dccps.nci.nih.gov/brp/presentations/slovic.pdf; accessed on July 10, 2006; Internet
Slutsky, E. (1915/1952) "On the theory of the budget of the consumer" In G. Stigler and K. Boulding (eds.) *AEA Reading in Price Theory*, Chicago: Richard Irwin Inc
Sober, E. (1984a) *The Nature of Selection: evolutionary theory in philosophical focus*, Cambridge: MIT Press
Sober, E. (1984b) *Conceptual Issues in Evolutionary Biology: an anthology*, Cambridge: MIT Press
Sober, E. and D. Wilson (1998) *Unto Others: the evolution and psychology of unselfish behavior*, Cambridge: Harvard University Press
Sousa, R. de (1987) *The Rationality of Emotion,* Cambridge: MIT Press

Sperber, D., F. Cara and D. Girotto (1995) "Relevance Theory Explains the Selection Task", *Cognition*, 57, pp. 31-95

Sperber, D. (1996) *Explaining Culture: a naturalist approach*, Oxford: Blackwell
Starmer, C (1999) "Experiments in Economics: should we trust the dismal scientists in white coats?" *Journal of Economic Methodology* 6, pp. 1-30
Starmer, C. (2000) "Developments in Non-Expected Utility Theory: the hunt for a descriptive theory of choice under risk", *Journal of Economic Literature*, 38, 2, June, pp. 332-382
Steedman, I. (1989) *From Exploitation to Altruism: aspects of political economy*, Boulder: Westview Press

Stein, E. (1996) *Without Good Reason*, Oxford: Clarendon Press
Stich (1983) *From Folk Psychology to Cognitive Science: the case against belief*, Cambridge: MIT Press

Stich, S. and Ravenscroft (1993) "What is Folk Psychology?" *Cognition*, 50, 1-3, pp. 447-468
Stigler, G. (1950) "The Development of Utility" Journal of Political Economy, 58, nos. 4 and 5, pp. 307-327; 373-396
Stigler, G. (1975) "Smith's Travels of the Ship of State," In A. Skinner (ed.) *Essays on Adam Smith*, Oxford: Clarendon Press
Stigler, G. (1986) "Ethics or Economics?" in G. Stigler, *The Essence of Stigler*, Stanford: Hoover Institute Press
Stigler, G. (1950a) "The Development o f Utility Theory I," *Journal of Political Economy*, vol. 58, no. 4, pp. 307-327
Stigler, G. (1950b) "The Development of Utility Theory II," *Journal of Political Economy*, vol. 58, no. 5, pp.373-396
Stone, V., S. Baron-Cohen, A. Young, A. Calder and J. Keane (2003) "Amygdala and Orbitofrontal Involvement in Theory of Mind", *Neuropsychologia*, 41, 2, pp. 209-220
Stone, V., L. Cosmides, J.Tooby, N. Kroll and R. Knight (2002) "Selective Impairment of Reasoning About Social Exchange in a Patient with Bilateral Limbic System Damage" *Proceedings of the National Academy of Sciences of the United States of America*, 99, 17, pp. 11531-11536

Straub, P. and J. Murnighan (1995) "An experimental investigation of ultimatums complete information, fairness, expectations and lowest acceptable offers" *Journal of Economic Behavior and Organization*, 27, pp. 345-64

Stuss, D., G. Gallup and M. Alexander (2001) "The Frontal Lobes are Necessary for Theory of Mind", *Brain*, vol. 124, no. 2, Feb., pp. 279-286
Sugden, R. (1982) "On the Economics of Philanthrophy", *Economic Journal*, 96, 366, June, pp. 341-350
Sugden, R. (1984) "Reciprocity: the supply of public goods through voluntary contributions", *Economic Journal*, 94, 376, Dec., pp. 72-87
Sugden, R. (1991) "Rational Choice: a survey of the contribution from economics and philosophy". *Economic Journal*, vol. 101, no. 407, Jul., pp. 751-785

Sugden, R. (2005) "Experiment, theory, world: a symposium on the role of experiments in economics," *Journal of Economic Methodology*, vol. 12, 2, pp. 177-184

Thaler, R (1985) "Mental Accounting and Consumer Choice" *Marketing Science* 4, summer, pp.199-214
Thaler, R. (1980) "Toward a Positive Theory of Consumer Choice", *Journal of Economic Behavior and Organization*, 1, pp. 39-60
Thaler, R. (1991) *Quasi Rational Economics,* New York: Russel Sage Foundation
Thaler, R. (1992) The Winner's Curse: anomalies and paradoxes of economic life. New York: The Free Press
Thaler, R. (2000) "From Homo Economicus to Homo Sapiens". *Journal of Economic Perspectives*, vol. 14, no.1, pp. 133-41
Tanney, J. (1995) "Why Reasons May Not be Causes," *Mind and Language*, vol. 10, no. 1/2, pp. 103-126
Todd, P. (2001) "Fast and Frugal Heuristics for Environmentally Bounded Minds", In G. Gigerenzer and R. Selten (eds.), *Bounded Rationality: the adaptive toolbox*. Cambridge, MA. MIT Press
Todd, P. and S. Bullock (1999) "Made to Measure: ecological rationality in structured environments". *Minds and Machines*, 9, pp. 497-541
Todd, P. and G. Gigerenzer (2000) "Précis of Simple Heuristics that Make Us Smart" *Behavioral and Brain Sciences*, 23, pp. 727-741

Todd, P. and G. Gigerenzer (2003) "Bounding Rationality to the World", *Journal of Economic Psychology*, 24, pp. 143-165
Tooby, J. and L. Cosmides (1990) "The past explains the present: Emotional adaptations and the structure of ancestral environment". *Ethology and Sociobiology,* 11, 375-424
Tooby, J. and L. Cosmides (1992) "Psychological Foundations of Culture," In J. Barkow, L. Cosmides and J. Tooby (eds.) *The Adapted Mind*, New York: Oxford University Press
Trivers, R. (1971) "The Evolution of Reciprocal Altruism", *Quarterly Review of Biology*, 46, pp. 35-57
Tversky, A. and D. Kahneman, D (1983) "Extensional vs. intuitive reasoning: the conjunction fallacy in probability judgment," *Psychological Review*, pp. 293-315

Tversky, A. and D. Kahnmenan (1974) "Judgment and Uncertainty: heuristics and biases", *Science*, 185, pp. 1124-1131

Tversky, A., P. Slovic and D. Kahneman (1990) "The Causes of Preference Reversal", *American Economic Review*, 80, 1, pp. 204-217
Uchitelle, L (2001) "Following the money, but also the mind: some economists call behavior a key". New York Times, February 11th.
Vanberg, V. (1993) "Rational Choice, Rule-Following and Institutions: an evolutionary perspective". In U. Mäki, B. Gustafsson and C. Knudsen (eds.) *Rationality, Institutions and Economic Methodology*, London: Routledge
Veblen, T. (1909) "The Limitations of Marginal Utility" *Journal of Political Economy*, 17, 9, Nov., pp. 620-636
Verbeek, B. (2002) *Instrumental Rationality and Moral Philosophy: an essay on the virtues of cooperation*. Dordrecht: Kluwer Publishers
Vickrey, W. (1945) "Measuring Marginal Utility By Reactions to Risk" Econometrica, 13, pp. 319-333
Vromen, J. (1995) *Economic Evolution: an inquiry into the foundations of new institutional economics*, London: Routledge
Vromen, J. (2001) "Ontological Commitments of Evolutionary Economics," In U. Mäki (ed.) *The Economic World View*. Cambridge: Cambridge University Press
Wang, X. (2001) "Bounded rationality of economic man: new frontiers in evolutionary psychology and bioeconomics", Journal of Bioeconomics, 3, pp. 84-89
Wason, P. (1966) "Reasoning", In B. Foss (ed.) *New Horizons in Psychology*, Harmondsworth: Penguin
Watson, J. (1930) *Behaviorism*, New York: Norton
Williams, G. (1966) *Adaptation and Natural Selection: a critique of some current evolutionary thought*. Princeton: Princeton University Press
Williams, G. and D. Williams (1957) "Natural Selection of Individually Harmful Social Adaptations among Sibs with Special Reference to Social Insects", *Evolution* 11, pp. 32-39
Wilson, E. (1975) *Sociobiology.* Cambridge: Harvard University Press

Wong, S. (1973) "'The F Twist' and the Methodology of Paul Samuelson," *The American Economic Review*, Vol. 63, no. 3, June, pp. 312-325

Wu, G., J. Zhang and R. Gonzalez (2004) "Decision under Risk," Working paper. [Online] available from http://gsbwww.uchicago.edu/fac/george.wu/research/decision_under_risk_long.pdf; accessed on July 10, 2006; Internet

Zamagni, S. (1995) *The Economics of Altruism*. Hants: Edward Elgar
Zuriff, G. (1985) *Behaviorism: a conceptual reconstruction*, New York: Columbia University Press

Dutch Summary/ Nederlandse Samenvatting

Dit proefschrift is een poging om de meerwaarde te onderzoeken van interdisciplinair onderzoek door gedragseconomen, psychologen en neurowetenschappers. Het uitgangspunt ervan is het inzicht dat, anders dan de meeste praktiserende economen, hedendaagse gedragsonderzoekers het als de taak van een echte wetenschappelijke verklaring schijnen te zien kennis te bieden van de oorzakelijke geschiedenis van een verschijnsel (d.w.z. de keten van gebeurtenissen te beschrijven die het *explanandum* produceren) en daarmee inzicht te bieden in de plaats die het heeft in de causale structuur van de wereld. Dit boek richt zich op het onderzoeken van een controversieel filosofisch vraagstuk - de verklaring van besluitvorming en de niet te verwaarlozen rol die emoties daarbij kunnen spelen.

In plaats van standaardtheorieën uit de wetenschapfilosofie (en de filosofie van de economie) te gebruiken, koos ik ervoor onderzoek te doen naar de feitelijke praktijk van een onderzoeksgebied dat gedragseconomie heet. Dit betreft een onderzoeksagenda die erkent dat de standaard economische keuzetheorie niet in staat was economisch relevante gedragspatronen te voorspellen en te verklaren. Het gaat daarbij om gedragspatronen binnen drie belangrijke domeinen van economische analyse, te weten intertemporele keuze, beslissing onder risicovolle omstandigheden en prosociale keuze.

Deze dissertatie bestaat uit drie delen. Deel 1 biedt de lezer een historisch overzicht van de opkomst en neergang van economische verklaringen, ontdaan van expliciete psychologische aannames en vooronderstellingen. Deel 2 gaat over enige methodologische vragen (naast andere filosofische vragen) die een leidraad bieden voor verklaringen van keuzegedrag die emoties (en andere cognities) serieus nemen. Het schetst een theoretisch model dat bedoelt een verklaring van keuzegedrag op twee niveaus te bieden. Deel 3 onderzoekt keuzemodellen die tot doel hebben te komen tot verbeterde verklaringen (en voorspellingen) van concreet gedrag binnen de drie domeinen van de economische analyse: intertemporele keuze, beslissing in risicovolle omstandigheden en prosociale (coöperatieve) keuze. Bovendien vindt er een analyse

plaats van de vraag of een verklaring van besluitvorming op twee niveaus de causale keten van emotionele en cognitieve processen, die keuzegedrag tot stand kunnen brengen, op het spoor kan komen (inclusief die gedragspatronen die anomalieën genoemd worden).

In dit boek breng ik naar voren dat gedragseconomen beweringen doen over aard en doel van hun theorievorming en verklaring van keuzegedrag die passen binnen de traditie van wetenschappelijk realisme. Ik opper dat gedragseconomen kunnen worden beschouwd as realisten met betrekking tot keuzemodellen/-theorieën, als zij tenminste verbeterde modellen presenteren van het complexe proces van besluitvorming door middel van manipulaties die bedoeld zijn om de belangrijkste factoren in de productie van het *explanandum* te isoleren van alle overige. Deze herziene verklaringen van keuzegedrag beogen de verklarende en voorspellende capaciteiten van de economische analyse te verbeteren.

De belangrijkste stelling van deze studie is dat emoties niet te verwaarlozen mentale elementen vormen bij het proces van besluitvorming. Zij moeten om die reden nadrukkelijk opgenomen worden in een verbeterde economische benadering van gedrag, ondanks het feit dat zij slecht te traceren zijn in de analytische modellen. Ik presenteer en bediscussieer empirische bevindingen die aangeven dat modellen van intertemporele keuzes, beslissingen in risicovolle omstandigheden en prosociale (coöperatieve) keuzes, die emotionele processen opnemen, concreet waargenomen keuzegedrag beter verklaren en voorspellen dan modellen die dat niet doen.

Ondanks het verbeterde verklarende en voorspellende vermogen van de bestaande keuzegedragmodellen moeten er nog hiaten worden opgevuld. Deze dissertatie vestigt aandacht op een specifiek hiaat, namelijk hoe emoties deel uitmaken van de causale keten van mentale processen die uit economisch oogpunt belangrijke patronen van keuzegedrag bepalen. Mijn uitgangspunt is dat economen expliciet het moeilijke probleem van causaliteit aan de orde zouden moeten stellen en dat zij er hard aan moeten werken om

met een verklaring te komen die het causale traject onthult waardoor daadwerkelijk keuzegedrag tot stand komt (Hausman 2001).

Geïnspireerd door Simons theoretische keuzemodel probeer ik een model te ontwerpen dat een beschrijving geeft van de 'besluitvormingsmachine', die twee analyselagen heeft, een psychologische (mentale) en een neurale (brein). Een van de belangrijke punten in mijn boek is dat een verklaring van keuzegedrag die een 'brein' niveau van analyse bevat het verklarend vermogen van de economische analyse verbetert, aangezien zo'n verklaring bijdraagt aan het beantwoorden van de vraag òf (en hoe) bepaalde mentale elementen en processen het daadwerkelijke keuzegedrag bepalen.

Dit boek bevat de volgende hoofdstukken. Hoofdstuk 1 introduceert de hoofdvragen en legt de argumentatiestrategie uit. Hoofdstuk 2 geeft een historisch overzicht van psychologische economische verklaringen. Ik beweer dat modellen waarin geen expliciete psychologische aannames verwerkt zijn, werden ontwikkeld in de 19^{e} eeuw in reactie op twee filosofische problemen die als obstakels voor theoretische vooruitgang werden gezien. Het ene is het probleem van de meetbaarheid van hedonistische nuttigheid (hedonic utility) en het andere is de moeilijkheid om een positieve economische (markt)gedragstheorie te ontwikkelen die gebonden is aan de controversiële leer van het psychologisch hedonisme. In aanvulling daarop geef ik een verklaring voor de hedendaagse terugkeer van gedragseconomische modellen door te refereren aan twee andere problemen: het ene betreft de voorspelling van anomalieën en het andere het vinden van een voldoende verklaring van hoe keuzegedrag (inclusief anomalieën) tot stand komt in de werkelijkheid.

Hoofdstuk 3 gaat kort in op enige conceptuele en metatheoretische vraagstukken die gedeeltelijk de trend verklaren om te komen tot een hervorming van economische keuze- en gedragsmodellen. Door conceptuele verheldering en het gebruik van voorbeelden uit de literatuur tracht ik aan te tonen dat het pleidooi van gedragseconomen voor meer realistische psychologische aannames past binnen de traditie van wetenschappelijk realisme.

Hoofdstuk 4 schetst een theoretisch model dat beoogt keuzegedrag te verklaren op het niveau van ‘de geest’ en dat van ‘het brein’. Ik draag verder argumenten aan voor de hoofdrol die emoties en cognities kunnen spelen in het proces van manifest keuzegedrag. In aanvulling daarop stel ik dat een verklaring op twee niveaus bijdraagt aan verdere voortgang in theorievorming.

In de hoofdstukken 5 t/m 7 zijn gedetailleerde studies te vinden die gaan over verklaringen van keuzegedrag binnen de drie al genoemde domeinen van economische analyse. Deze hoofdstukken dienen twee hoofddoelen. Het ene is de hypothese te bevestigen dat bepaalde keuzegedragmodellen, door emoties serieus te nemen, meer en betere verklaringen van daadwerkelijk keuzegedrag kunnen bieden dan de conventionele nutstheorie. Het tweede doel is aan te tonen dat een theoretisch keuzegedragmodel met elementen uit neurowetenschap en evolutionaire psychologie goed overeenkomt met (empirische) modellen die de emotionele en cognitieve processen beschrijven die significant zijn voor economisch relevante patronen van gedrag.

Hoofdstuk 8 heeft als doel de argumentatie samen te vatten en kort sommige mogelijke implicaties van de voorgestelde dubbellaagse economische verklaring van keuzegedrag aan de orde te stellen. Tenslotte worden een aantal vragen geopperd ten behoeve van een verdere onderzoeksagenda.

Curriculum Vitae

Roberta Muramatsu was born in São Caetano do Sul, São Paulo, on November 17, 1973. She received her BSc and MPhil degrees in Economics from the University of São Paulo, Brazil. In 1996, her monograph *Neither the Wolf, nor the Nobleman, the place of ethically committed behavior in the economic world* won The Alfred Marshall Prize by the Liberal Institute of Brazil and the FIPE-University of São Paulo Best Monograph Prize. Her master thesis was entitled *On the possibility of altruism: an inquiry into the economics, psychology and biology of cooperative behavior* (1999). She was an assistant lecturer of economics from 1998 to 1999 in University of São Paulo. From September 1999 to June 2000 she started the MPhil program in Erasmus University Rotterdam at the Erasmus Institute for Philosophy and Economics.

In 2003 she became a lecturer of Economics at Ibmec Business School in São Paulo. Since 2004 she is also a lecturer of economic methodology and history of economic thought at Mackenzie University of São Paulo. Her latest published articles are entitled "Emotions as Mechanisms for Boundedly Rational Agents: the fast and frugal way" (with Yaniv Hanoch), *Journal of Economic Psychology*, April 2005, 26, pp. 201-221, and "Ulysses' return: notes about the rationality of commitments and plans" (with Ana Maria Bianchi), *Revista de Economia Politica*, vol.28, no. 2, April-June, pp. 23-44