THREE CORRELATED ESSAYS ON THE NEURAL FOUNDATIONS OF ECONOMIC DECISION-MAKING

Thesis by

Ming Hsu

In Partial Fulfillment of the Requirements for the

degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2006

(Defended July 31, 2006)

ACKNOWLEDGEMENTS

My introduction to neuroeconomics came at a seminar at the University of Arizona when I was an undergraduate. Kevin McCabe was presenting his paper on neuroimaging of the Trust game. A little light bulb went off in my head and I sort of thought to myself how special it could be. Little did I know at the time that this was the first study in the now booming, albeit still small, field of "neuroeconomics". I spent a year at Economic Science Laboratory, and during that time, Kevin, Vernon, and Dan Houser were instrumental in guiding me in my choice of research, and my choice of Caltech as my graduate school.

My time at Caltech has been immensely satisfying at both the intellectual and the personal level. Caltech has been a great research environment. I can't imagine another place that would have given me so much freedom and had such accessibility to the faculty. The people with whom I regularly work with include not only economists and neuroscientists, but also psychologists, philosophers, and even electrical engineers (!).

Peter Bossaerts, Matthew Jackson, Charles Plott, Steve Quartz have been particularly instrumental for their intellectual input. I am also grateful for the wise words of Ken Binmore, for his unwavering belief that the truth is not an abstraction or inconvenience that economists have to but pay lip service to.

My greatest debt for this thesis and my intellectual progression is certainly to my advisor and mentor, Colin Camerer. Aside from never being short on Britney or Paris stories, he was a constant source of motivation and ideas. Perhaps, as is common with graduate students, I am long on procrastination and loathe letting go. I shudder to imagine what would have happened without someone to backward induct for me the consequences of my actions in my graduate career.

I would like to give a big shoutout to all my wonderful friends at Caltech and Southern California, especially Cedric Anen, Meghana Bhatt, John Eifler, Michelle Feng, Vanessa Heckman, and Brian Zid. I have grown to love Los Angeles, with its wonderful food, music, nightlife, and people. I don't think I would have enjoyed myself these five years nearly so much without all of you. It is with great regret that I leave you guys.

The economics job market is a stressful for most people under any circumstances. Being the first one in this new field with the catchy and perhaps strange-sounding name (neuroeconomics) makes it even more stressful. I was very grateful for the warm reception given me and my work by everyone who interviewed me at Allied Social Science Association 2006, especially Soo Hong Chew, Chris Foote, Tanjim Hossain, Bart Taub, and Steve Williams. It could have been a desolate and depressing experience, but it was not so, thanks to all of you.

Finally, I would like to reserve my last words for my family. Chinese parents get a bad rep in the U.S. They are alternately known as demanding, overbearing, and closed-minded—some of it well deserved. I consider myself fortunate to have parents who are none of these. It is with great fondness to recount my family's support in my career, even (especially) the many instances when they smiled wistfully at, what must have been for them, the strange idea of 脑袋经济.

ABSTRACT

Paradoxes are useful in science because they hint at errors or inconsistencies in our models of the world. In this thesis, I study two well-known and long-standing paradoxes in decision theory from the point of view of neuroeconomics. This approach combines tools and ideas from economics and neuroscience, and tries to understand the neural mechanisms and the causal structures behind these paradoxes.

Since its introduction in Ellsberg (1961), the Ellsberg Paradox has been one of the most studied violations of subjective expected utility theory (SEU). In one version of the paradox, a decision-maker is confronted containing two urns with 100 balls that are either red or blue. In the first (risky) urn, she is told there are 50 red and 50 blue; whereas no further information is given about the second (ambiguous) urn. A commonly observed choice pattern is for decision makers to choose to bet on both red and blue in the first urn. Clearly, if probabilities are additive, such rankings are inconsistent with SEU.

First, I present brain imaging that shows that the brain treats risky and ambiguous choices differently. This is done through the use of functional magnetic resonance imaging (fMRI), a method that measures brain activity indirectly through blood flow. I find evidence that the brain regions respond differently to ambiguity and risk. Furthermore, the region that is correlated with expected money value of choices are activated more under risk than ambiguity, confirming that expected utility of ambiguous gambles are lower than those of equivalent risk gambles. Finally, the temporal relationship between the regions suggests a network where one brain region signals the level of uncertainty (amygdala), sent through another region (orbitofrontal cortex), and increases (decreases) expected utility of the choices, represented in the activity of a third region (striatum).

Brain imaging results, however, is limited by its correlational nature. To assess necessity, if a particular brain region causes a certain behavior, taking it out should remove that

behavior. Conversely, to assess sufficiency, stimulating the brain region should create that behavior.

In the former, I study patients who have damage to the orbitofrontal cortex (same region found in the brain scans). I find that these patients were both ambiguity- and risk-neutral. This compares to ambiguity- and risk-averse behavior of patients with damage to other parts of the brain not found in the brain scans, similar to normal individuals. This confirms the idea that specific brain regions are necessary for distinguishing between risk and ambiguity. In the latter, I activate amygdala of (normal) subjects through mild electrical stimulation (a method known to elicit activation of the region). This allows us to test whether this method increases the ambiguity/risk aversion of subjects.

The third chapter studies the Allais Paradox and the probability weighting function. The fact that people do not appear to weight probabilities linearly as dictated by subjective expected utility theory has been known since the 1950s. More specifically, people have been found to overweight small probabilities, and underweight large probabilities. This chapter has two goals. First, I attempt to find the neural correlate of the probability weighting function: that is, is the probability weighting function as discussed in the decision theory literature found in the brain? Second, I posit a hypothesis for the generation of the probability weighting function with data from psychophysics and neuroscience. Together they shed light on how the brain encodes probabilities as a physical quantity as well as how it might combine decision weights and rewards to calculate expected utility.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

This thesis is about the emerging field of neuroeconomics. Neuroeconomics seeks to ground theories and models of economic behavior in terms of its underlying neural mechanisms. Recent advances in technology have allowed us to study economic phenomenon, such as decision theory and game theory, much as neuroscientists in the past century did with vision, audition, or olfaction.

These advancements have impacted a number of fields, spanning a diverse range of questions and methodologies—ranging from neurophysiology, psychophysics, to cognitive neuroscience. Some of the most successful applications of neuroscientific data have occurred in the field of psychophysics, in the realm of some fundamental perceptual apparatus of organisms.

Psychophysics is the study of the relationship between physical stimuli such as light, sound, odor, and their perception. The assumptions underlying the neuroscientific studies of psychophysics read very much like those one would imagine a neuroscientific study of economics would read. For example, Boynton et al. (1999) stated the following set of assumptions using functional magnetic resonance imaging (fMRI) to test psychophysics:

1. fMRI responses proportional to local average neuronal activity.

2. Subjects' perception result of neuronal responses.

3. Variability in psychophysical measurements due to additive (IID) noise.

Although Boynton et al. stated "perception", to operationalize (2), it is necessary to make the assumption that perception translates accurately to behavior. Whereas, in

psychophysics, the neuronal responses are to stimuli such as luminosity, contrast, or decibel levels, in neuroeconomics, they include variables such as probabilities, expected value, or constructs such as risk and ambiguity.

1. Both behavior and perception are the result of neuronal responses.

2. Our method of measurement adequately captures the neuronal responses.

The first assumption is the central one in any study involving neuroscience. It states that the phenomenon of interest is the result of the collective activity of the neurons of the nervous system. For psychophysicists, the phenomenon is the perception of physical stimuli, whereas for economists, it can range from perception of abstract constructs such as probabilities, or choice behavior.

In the case of psychophysics, this assumption is well rooted in the neurophysiological literature. For example, one of the classic papers on the visual system by Hubel and Wiesel (1959) showed that neurons in the visual cortex of the cat selectively respond to lines oriented at specific angles in the visual field (see Figure 1). Still other neurons selectively respond to lines at certain angles moving in one direction. They called these "simple" and "complex" cells, respectively.

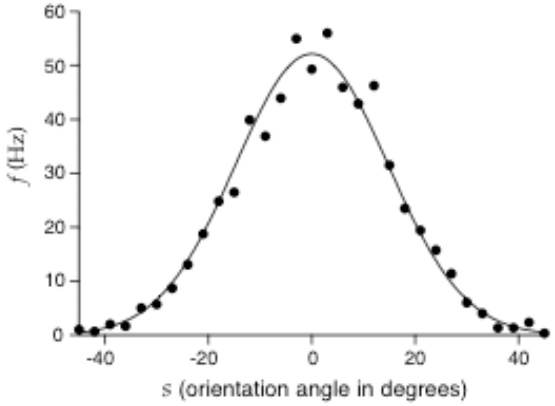


Figure 1: Average firing rate and fitted (normal) distribution of a cat V1 neuron plotted as a function of orientation angle of a bar of light. Adapted from Dayan and Abbott (2001).

There are a number of other ways that a neuron can encode information. The simplest, and most widespread, is the intensity code, or rate code. It has been known since the 1950s that the firing rate of neurons encode quantities such as luminosity and contrast. First discovered by Lord Edgar Adrian, this is the phenomenon whereby the rate of the neuronal firings is proportional to the intensity of the stimuli (Rieke 1997). Much of this seminal work was done on the *Limulus* (horseshoe crab), by recording from its compound eye (Figure 2).

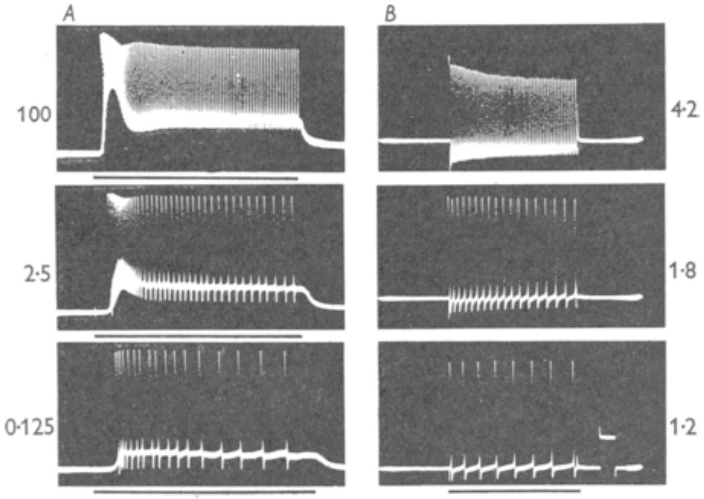


Figure 2: Responses to light and to electrical current. (A) Response to light: numbers on left indicates relative intensity; duration of illumination is indicated by solid line below each graph. (B) Response to current: numbers on right indicates intensity of depolarization in nA. Time line at bottom right, 1sec. (adapted from Fuortes et al. (1959)).

Statistical Process of Neuronal Firing

The stochastic process by which neuronal firings are generated is typically modeled by a Poisson process.

$$P_T(n) = \lim_{\Delta t \to 0} \binom{M}{n} (r\Delta t)^n (1 - r\Delta t)^{M-n}$$

$$P_T(n) = \frac{(rT)^n}{n!} \exp(-rT).$$

One of the properties of the Poisson distribution is that the mean equals the variance. The ratio between these two quantities, $\sigma_n^2 / E(n)$, is called the *Fano factor*, and is a homogeneous Poisson process. Experiments estimating this variation through fitting via a

power function $\sigma_n^2 = AE(n)^B$ have determined that both A and B to be between 1.0 and 1.5. Figure 3 shows a plot of this relationship from (Gur et al. 1997).

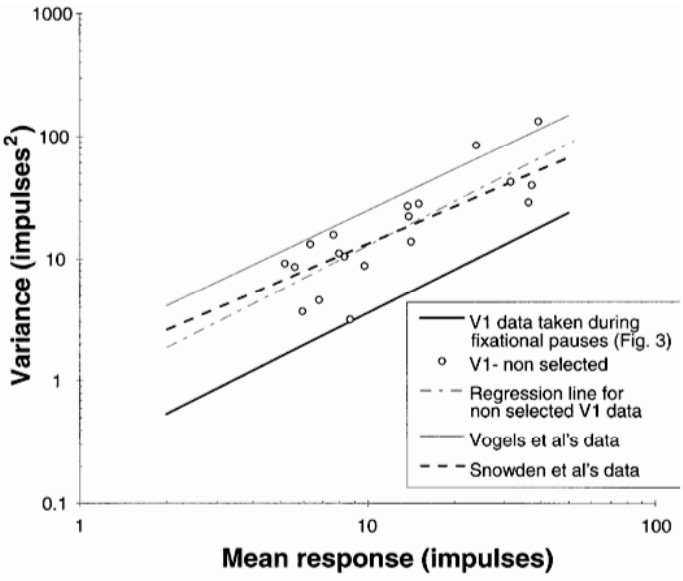


Figure 3: Relationship between response variance and response strength in the monkey visual cortex (adapted from Gur et al. (1997).

Neuroeconomics in Historical Context

Clearly, the approach taken in this thesis, and neuroeconomics in general, is a sharp turn in economic thought. Around the turn of the century, economists made a clear methodological choice to treat the mind as a black box and ignore its details for the purpose of economic theory. In an 1897 letter Pareto wrote:

> *"It is an empirical fact that the natural sciences have progressed only when they have taken secondary principles as their point of departure, instead of trying to discover the essence of things. ...Pure political economy has therefore a great interest in relying as little as possible on the domain of psychology" (quoted in Brusino 1964 p. xxiv).*

Pareto's view that psychology should be ignored was reflective of a pessimism of his time about the ability to ever understand the brain.[1] As William Jevons (1970) wrote. prior to Pareto, in 1871:

> *"I hesitate to say that men will ever have the means of measuring directly the feelings of the human heart. It is from the quantitative effects of the feelings that we must estimate their comparative amounts."*

This pessimism about understanding the brain, in part, led to the popularity of "as if" rational choice models. Models of this sort posit individual behavior that is consistent with logical principles, but do not put any evidentiary weight on direct tests of whether those principles are followed. For example, a consumer might act as if she attaches numerical utilities to bundles of goods and choose the bundle with the highest utility, but if asked to assign numbers directly, her expressed utilities may not obey axioms like transitivity. The strong form of the as-if approach simply dismisses such direct evidence as irrelevant because predictions can be right even when they are based on incorrect assumptions (e.g., Friedman 1964).[2]

As-if models work well in many respects, but tests of the predictions that follow from as-if rational choice (as well as direct tests of axioms), have also established many empirical anomalies. Behavioral economics describes these regularities and suggests formal models to explain them (e.g., Camerer 2003). Debates between rational-choice and behavioral models usually revolve around psychological constructs, such as loss-aversion (Kahneman and Tversky 1979) and a preference for immediate rewards, which have not been observed directly. But technology now allows us to open the black box of the mind and observe brain activity directly. The use of data like these to constrain and inspire economic theories, and make sense of empirical anomalies, is called "neuroeconomics" (Zak 2004; Camerer et al. 2005; Chorvat and Mccabe 2005).

Criticisms of Neuroeconomics

The leap from neurons to behavior is a large one. Perhaps inevitably, it has attracted criticism. Gul and Pensendorder (2005), in particular, makes the vociferous argument that economists should not, on principle, pay attention to the details of the brain.

One of their central arguments is that the idea of utility underlying revealed preference theory is strictly *as-if*—"Standard economics does not address mental processes and, as a result, economic abstractions are typically not appropriate for describing them" (Gul and Pesendorfer 2005). Therefore, according to the authors, any attempts by neuroeconomists to try to find the neural basis of utility is neither here nor there, since it doesn't exist in the first place. The spirit of this argument can be traced back to Milton Friedman's famous dictum that economic theories should be judged by the accuracy of their predictions rather than the plausibility of their axioms (Friedman 1964). Friedman's view was that (1) theories should be judged by accuracy of predictions; and (2) false assumptions could lead to accurate predictions.

The central assumption of neuroeconomics, and of this thesis, is that creating more realistic assumptions will lead to better predictions. Furthermore, a critical examination of assumptions is healthy and critical to the advancement of any science. In this regard, neuroeconomics shares the emphasis on accuracy in principle (1), but also bets on the possibility that improving the accuracy of assumptions will lead to more accurate predictions.

An analogy to organizational economics illustrates the potential of neuroeconomics. Until the 1970's, the "theory of the firm" was basically a reduced-form model of how capital and labor are combined to create a production function, as the basis for an industry supply curve. Contract theory opened up the blackbox of the firm and modeled the details of the nexus of contracts between shareholders, workers and managers (is the three elements of the firm). The new theory of the firm replaces the (still useful) fictional profit-maximizing firm which has a single goal, with a more detailed account of how components of the firm interact and communicate to determine firm behavior. Neuroeconomics proposes to do the

same by treating an agent like a firm: Replace the (useful) fiction of a utility-maximizing agent who has a single goal, with a more detailed account of how components of the agent's brain interact and communicate to determine agent behavior.

The success of the rational actor model in "as if" applications shows that this level of detail is not necessary for certain sorts of analysis, especially those that deal with populations of decision makers instead of individuals. For example, neuroeconomics will never displace the powerful concepts of supply and demand, and market equilibrium. However, a deeper understanding of the mechanics of decision-making will help us understand deviations from the rational model better. Knowing the process of decision making should allow us to understand not only the limits of our abilities to calculate optimal decisions, but also the heuristics that we use to overcome these limits.

Furthermore, in most areas of behavioral economics there is more than one alternative theory. Often there are many theories that are conceptually different but difficult to separate with current data. To the extent that some of these theories commit to neural interpretations, the brain evidence can help sort out which theories are on the right track and also suggest new theories.

Similar Criticisms of Cognitive Neuroscience

Much of the criticisms raised by Gul and Pesendorfer are similar in spirit to those raised by psychologists and philosophers decades ago (Churchland 1986). As Dehaene (2003) recounts of the progress of cognitive neuroscience

> *"Today, however, we know that this view was unnecessarily narrow. The new cognitive neuroscience routinely mixes psychological and neural observations in the same experiments. Psychological concepts are not ruthless eliminated, as was initially foreseen by the most opinionated anti-functionalist philosophers... Rather, they are enriched, constrained and transformed by the accruing neural data."*

The context of Dehaene's quote surrounds the discovery of neurons that represent numbers via logarithmic encoding, thus shedding light on an age-old debate of whether the neural basis of the Weber-Fechner law (the finding that the threshold of discrimination between two stimuli increases linearly with stimulus intensity), and the postulation that internal representation of the stimuli follows the logarithm of the external stimuli.

Whereas, the interplay between the neural and the psychological is now relatively uncontroversial, skepticism over use of functional imaging techniques can be more vocal, perhaps owing in large part to the still youthful state of the field (fMRI was developed in the early 1990s, whereas Weber and Fechner lived in the 19th Century). Concern over the use and abuse of fMRI was especially prominent in social psychology in the late 1990s and early part of the new millennium (see e.g. Shulman 1996; Hardcastle and Stewart 2002; Willingham and Dunn 2003).

Shulman, in particular, outlined four assumptions that are used by cognitive psychologists in the interpretation of fMRI data, two of which he finds objectionable.

1. Differences between cognitive tasks map to unique regions of the brain or activate the same region differentially;

2. Tasks can be designed to have unique (non-shared) cognitive processes;

3. The difference images of two tasks (either by fMRI or PET) will reveal the location of the unique (non-shared) cognitive components;

4. The statistical validity of the difference image can be established quantitatively.

Shulman proposes that although (1) and (4) have firm basis in the neuroscience literature, (2) and (3) are problematic. Much of the criticism centered on the lack of specific and parsimonious hypotheses. This can be seen by the phrasing of (3), as most of the early studies used a technique called "cognitive subtraction", which are essentially t-tests between designated "experimental" condition and "control" condition.

Although a naïve implementation of this technique clearly has methodological limitations, there are at least two strong arguments against this critique. First, perhaps more so than social and cognitive psychology, economics possesses well-established models that make precise, quantitative predictions. More importantly, many of these models, such as the expected utility model, are very specific about the choice variables, as well as having a corpus of econometric techniques to test these models.

Second, and more importantly, modern neuroscience possesses a wide variety of tools that, together, create complementarities and allow each to compensate the weaknesses of the other. We review some of these tools in the next section.

Tools of Measuring the Brain

Much of the potential of neuroeconomics comes from recent improvements in technology for measuring the brain. There are now over 10 established techniques in neuroscience, each with its own strengths and weaknesses. Some common tradeoffs include spatial and temporal resolution, invasiveness, and ability to assess direction of causality. Figure 4 depicts the characteristics of these methods in terms of their spatial and temporal resolution, as well as their invasiveness.

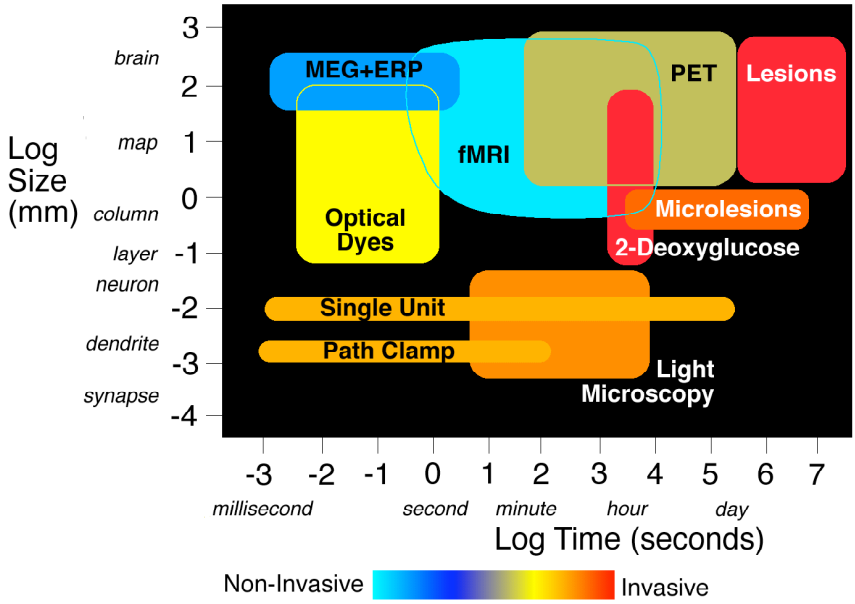


Figure 4: Temporal versus spatial resolution of various methods of studying the brain. Heat map is the invasiveness of the method. Abbreviations: MEG=magneto-encephalography; ERP=evoked response potentials; *f*MRI=functional magnetic resonance imaging; PET=positron emission tomography. (Adapted from Churchland and Sejnowski 1988).

Of these tools, functional magnetic resonance imaging (fMRI) is perhaps the most popular today, especially when working with human subjects. Its popularity is derived from two main sources. First, it is non-invasive. Second, fMRI provides relatively good temporal and spatial resolution. Images are taken every few seconds and data are recorded from voxels which are 3-4 cubic millimeters in size.

Better spatial and temporal resolution is available in single unit or multiple unit recordings, which uses electrodes to record brain activity at the neuron level. Because of the need to implant electrodes directly into the brain, however, these studies are limited mostly to animals, including nonhuman primates, and to some neurosurgical patients.

Older tools like electroencephalogram (EEG), recording very rapid electrical activity from outer brain areas, and psychophysiological recording (skin conductance and pupil dilation, for example), continue to be useful, often as complements that corroborate interpretations

from other methods. For example, optical dyes can be used to measure the membrane potential of the neuron—but due to its invasiveness, is not used in human research.

Functional neuroimaging is especially powerful when combined with existing techniques. One venerable method is the study of animals and patients with localized lesions to specific areas of the brain. Lesion studies provided some of the earliest evidence toward the link between brain and behavior, and are still invaluable in assessing the often complex causal neural networks involved in a given behavior.

Lesion patients can show startling dissociations between tasks that seem remarkably similar. For example, there are lesion patients who can write well, but can't read, a condition referred to as *alexia without agraphia* (Caffarra 1987); patients with *anterograde amnesia*—the inability to form new memories—who do not form conscious declarative explicit memories for emotions or physical procedures (Scoville and Milner 1957). However, these patients can form memories implicitly; that is, they know information but don't know that they know it. This dissociation indicates that declarative memories, and meta-knowledge are stored separately.

Finally, with the use of psychopharmacological drugs, transcranial-magnetic stimulation (TMS), and deep brain stimulation, it is also possible to directly alter brain activity. This helps to establish whether certain brain processes are *sufficient* for some type of behavior. For some examples, see Kosfeld et al. (2005), who found that the administration of the hormone oxytocin increases the level of repayment in the Trust game. For TMS, see Zangaladze et al. (1999), who were able to disrupt tactile discrimination of grating patterns by applying TMS either on the visual cortex or the somatosensory cortex. The practical limitation of using TMS is that only surface cortical areas can be reached. In the case of psychopharmacological agents, their effects on the brain tend to be diffuse, which makes it difficult to pinpoint the specific mechanism at work.

Fundamentals of fMRI

Because of its extended use in this thesis, its complexity, and its relative unfamiliarity to most economists, I will spend this section describing the basic aspects of fMRI and fMRI data.

Physics of fMRI

Comprehensive guides to the physics of both MRI and fMRI can be found in Hornak (1997) and Huettel, Song, and McCarthy (2004)|.

MRI (and fMRI as a consequence) takes advantage of the net spin of atomic nuclei. This, combined with the fact that different atoms exhibit different spin properties, means that distinct tissues throughout the body also have distinct spin properties. Moreover, these properties may change over time; in the case of fMRI, this is brought about by local changes in oxygenation in the blood flow.

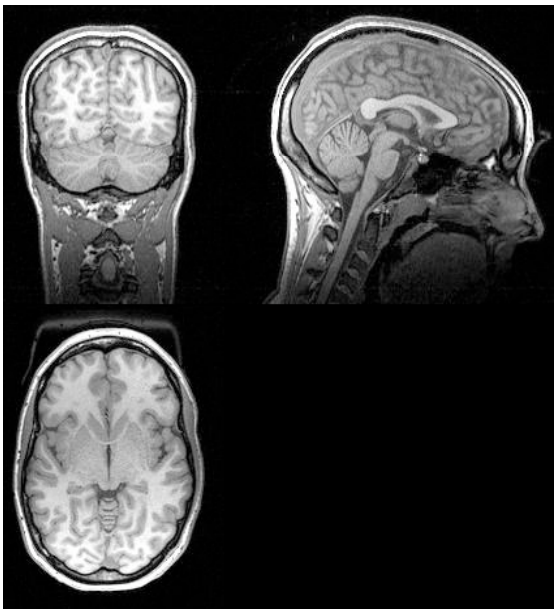


Figure 5: Example of (static) MRI anatomical image. Clockwise from top left: coronal slice (viewed from back of brain), sagital slice (viewed from left side), and axial slice (viewed from top).

Specifically, the magnetic properties of hemoglobin change depending on its oxygenation. Oxygenated hemoglobin is diamagnetic and has similar magnetic properties as the rest of the brain. However, deoxygenated hemoglobin is paramagnetic. This causes a change in the magnetic susceptibility of the local blood supply when the metabolic demands of the

tissue deoxygenates nearby hemoglobin. Hence fMRI measures what is called blood oxygenation-level dependent (BOLD) signal. Ogawa et al. (1990) first showed that when mice breathed from different concentrations of oxygen, the low concentration oxygen resulted in significant signal loss in blood vessels within the brain. In the decade ensuing, tremendous progress was made in enabling the measurement of BOLD activity that is available today.

Correlation between fMRI and Neuronal Activity

This is the fundamental question in the use of fMRI to study neural correlates of behavior. Although the correlation between fMRI and neural activity has been known for over 100 years (Roy and Sherrington 1890), the precise relationship is surprisingly complex and is still under active research. Here, we review some of the current evidence on the nature and strength of this relationship.

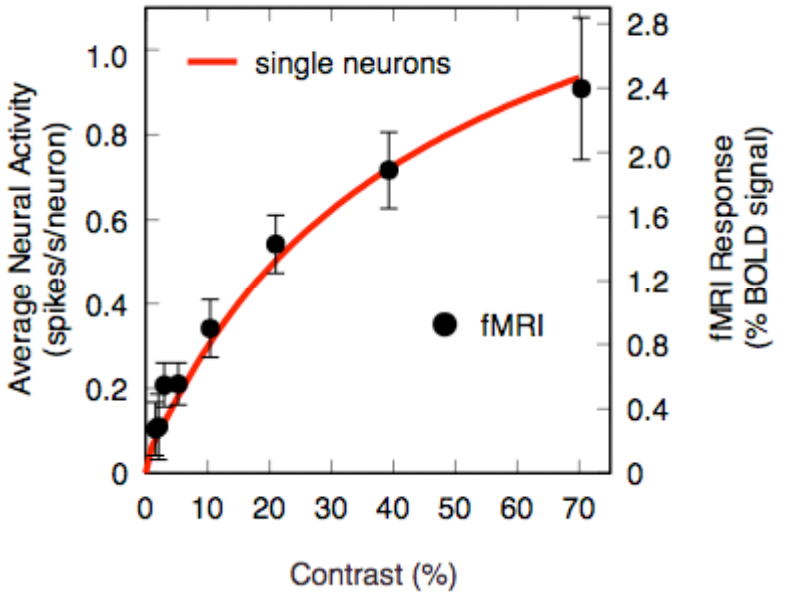


Figure 6: fMRI response in human V1 and average firing rate of monkey V1 (*adapted from Heeger et al. (2000)*).

Rees et al. (2000) and Heeger et al. (2000) used measurements of BOLD via fMRI with humans to find functional homologues to electrophysiological recordings in monkeys. Heeger et al. estimated the average neuronal response to the visual perception of contrast in

the monkey visual cortex V1 were proportional to that of the BOLD response in the human V1 (Figure 6).

Although suggestive, both Rees et al. and Heeger et al. were limited to correlating the measurements between two *species*. Logothetis et al. (2001), in a heroic effort, simultaneously recorded multi-unit activity (MUA), local field potential (LFP), and BOLD response from the monkey visual cortex. Figure 7 shows the distribution of r^2 values between LFP and BOLD as well as MUA and BOLD. The distributions possess heavy negative skewness. The mean r^2 value of MUA and BOLD is 0.521, whereas the mode is 0.672. Similarly, the mean r^2 value of LFP and BOLD is 0.445, whereas the mode is 0.457.

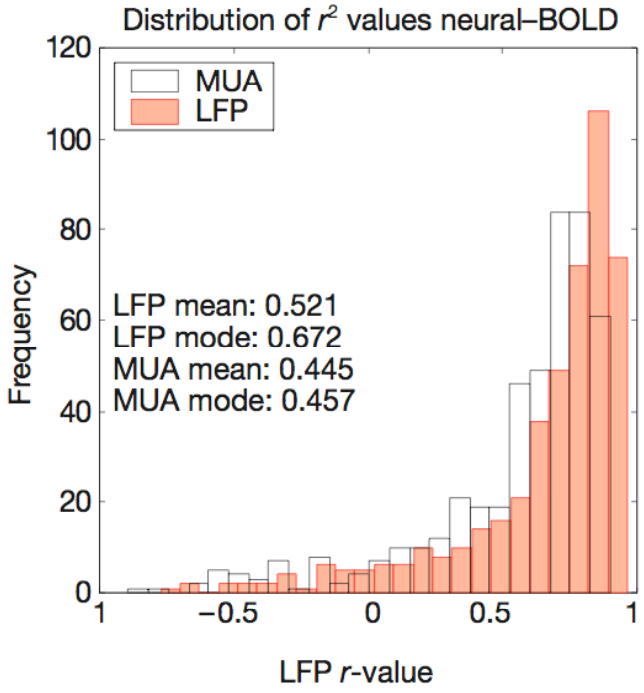


Figure 7: Correlation of neural activity and BOLD response (adapted from Logothesis et al. (2001)).

Ress, Backus, and Heeger (2000) was able to link fMRI response of the visual cortex and perception of contrast. Subjects viewed a uniform gray background, and were asked to indicate whether a plaid pattern was present briefly. Figure 8 shows the correlation between fMRI response and performance in the threshold detection task.

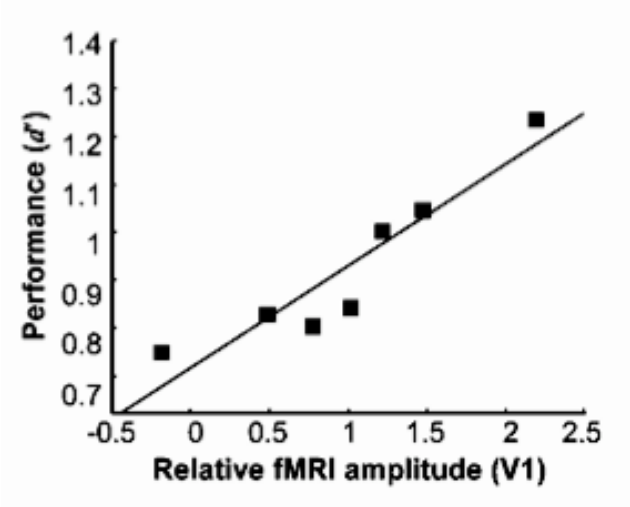


Figure 8: Scatterplot of behavioral performance in threshold detection task (y-axis) versus fMRI activity in V1 (x-axis) and linear regression line (r=0.92, p<0.001). (adapted from Rees, Backus, and Heeger (2000)).

fMRI Data

Here we discuss briefly the nature of fMRI data. As in previous sections, no attempt at comprehensiveness is made. Focus is placed on the aspects of fMRI data that are likely to be novel to economists and econometricians.

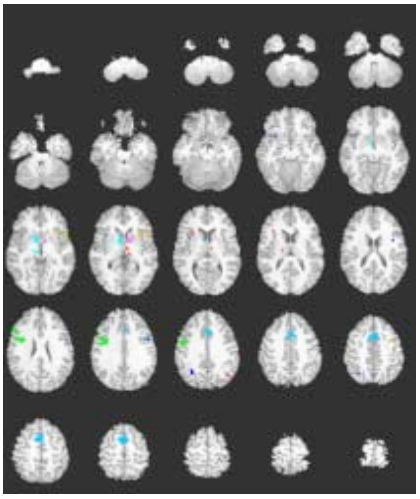


Figure 9: Sample cross-section of fMRI data. Each panel represents one 2-D slice of the brain. fMRI scanners acquire these in-plane slices sequentially. Colors indicate significantly activated voxels.

MR scanners encode the data in frequency domain, which are then processed into the spatial domain. Usually, two sets of images are obtained—a set of high-resolution anatomical images (Figure 5), and a set of lower-resolution functional images during the course of the experiment (Figure 9). As the anatomical images are used only for localizing

the brain regions and aid in normalizing the data between subjects, the term "fMRI data" generally refers to the latter, functional images.

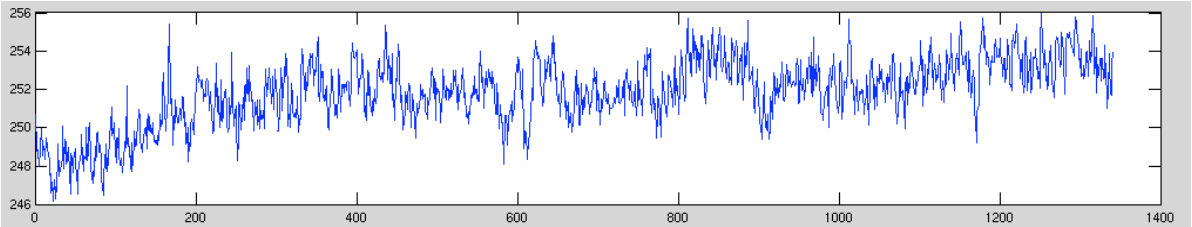


Figure 10: Sample time series of fMRI data taken from a voxel within the brain.

Functional MR data constitute a 3D panel dataset. The MR scanner takes in-plane snapshots of the brain (2D) and proceeds to cover the brain iteratively on the order of seconds to obtain the time series (Figure 10). The time that it takes for the scanner to acquire one full cycle of data through the brain is called the TR, and in our case it is 2 seconds.

Main Sources of Noise in fMRI Data

One limitation of fMRI is that its signal-to-noise ratio is generally low, on the order of 0.25-0.5% (Huettel et al.). A wide variety of different sources contribute to the overall noise, from the scanner to the subject. Much of the effort in the design, running, and statistical analysis of fMRI data consists of minimizing and taking account of these noise sources in a principled manner. Here, we discuss some of the main sources of noise.

Thermal Noise. These are (white) noise created by the electronic circuits of the scanner. This is perhaps the biggest contribution of noise to fMRI data.

Scanner Drift. This is caused by small instability in the scanner gradients. Typically, this introduces linear trend or low frequency noise to the data.

Subject Motion. Perhaps the most serious source of noise, subject motion creates partial voluming effects. This can be significant as signal intensity of nearby voxels may differ

substantially, especially in voxels near the boundary of tissues (such as around the ventricles, which contain cerebral spinal fluid, or the edges of the brain). Much effort is taken, therefore, to stabilize the subject during the course of the experiment. The remaining motion is corrected using motion correction algorithm described in the section below.

Physiological Artifacts. This includes blood flow changes induced by the cardiac and respiratory cycles. These typically occur at a higher temporal frequency than signals of interest. They can either be filtered out with a lowpass filter or estimated as part of the autocorrelation function. The latter is currently preferred.

fMRI Data Preprocessing

The statistical package SPM (Statistical Parametric Mapping, version SPM2), was used for the fMRI data in all the studies of this thesis (Friston et al. 1995b; Friston et al. 1995c). A detailed introduction to computational neuroanatomy and statistical parametric mapping can be found in the online book Human Brain Function by Ashburner, Friston, and Penny (2001). Preprocessing is done in the order described in the following section.

Slice timing correction. MR images are acquired through iterative snapshots through the brain. Because of the sequential nature of the process, the time series are not matched throughout the brain. Following Friston et al. (1995a), we align the time series within the brain using a sinc weighting interpolation function.

Realignment. As mentioned before, head motion is one of the major sources of artifacts of fMRI data. This arises from the assumption that the head is stationary throughout the data acquisition process. Two complementary methods are used to minimize this problem, one during data collection, and one during data analysis. At the data collection stage, head restraints are used to minimize possible movement during the experiment. This ensures that the movements are small (maximal drift is typically less than ½ the dimension of a voxel, or ~1.5mm, e.g., Figure 11). The residual movements are corrected for via a realignment process using rigid body transformations (Friston et al. 1995a). Note that

because all motion correction is done within subject, and that the shape of the brain does not differ appreciably over the course of the experiment, the transformations essentially reduce to those of translation and rotation in space.

Finally, because of the computational requirements to estimate the movement correction coefficients over many thousands of images potentially, SPM2 employs a "local" approach in which the images are realigned using an initialization point of the previous image, and then minimizing the mismatch between the images by adjusting the images in an iterative manner. This procedure can potentially have local maxima problems, but given the computational savings, and the fact that the physical restraint limits the amount of movement, this is a reasonable approach and is widely used.

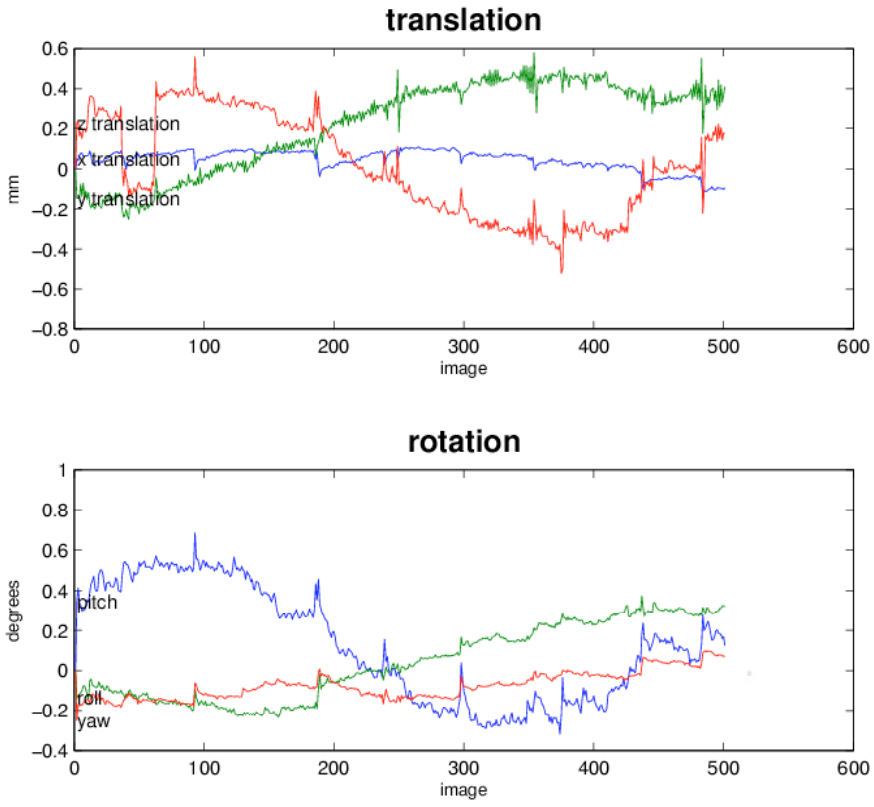


Figure 11: Typical estimated motion parameters from SPM2. The x-axis of the panels denote time, also known as images.

Coregistration. This step refers to mapping the low-resolution function MRI images to the high-resolution T1 image. This step may be considered optional, as all functional images will be standardized to a common space in the next step anyway. However, because coregistration affords some benefits to this standardization process, we employ it here. The algorithm uses an information-theoretic approach—namely maximizing the mutual information of the joint histogram of the functional images and the T1 (Friston et al. 1995a). Mutual information maximization is has the advantage that it can take into account nonlinear dependencies of the data, as opposed to only linear ones as in the case of Pearson's correlation. This is helpful as images from different modalities differ greatly in the intensity scales, and often with nonlinear relationships. For this reason, mutual information enjoys widespread use in medical imaging.

Normalization. In order to generalize findings and make inferences across individuals, one must take into account the variation in the brain structure between people, which can be significant. This is achieved through a process called normalization. SPM2 uses both linear and non-linear warping techniques to normalize the data to a standard canonical brain (Ashburner and Friston 1999). This allows one to calculate both group statistics as well as standardizing the reporting of results between different research groups and paradigms.

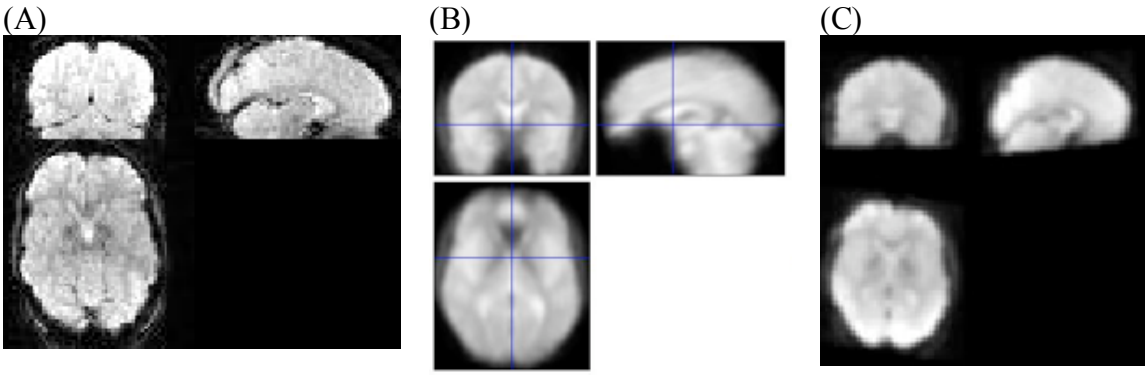


Figure 12: (A) Original un-normalized brain, (B) Standard canonical, or template, brain. (C) Normalized brain from (A).

Smoothing. Because of the relatively poor signal-to-noise ratio (SNR) of fMRI time series data, it is typical to spatially smooth the data. This is done by convolving the data with a Gaussian kernel parameterized by the full-width-at-half-maxima (FWHM). The choice of FWHM depends upon the underlying (unknown) SNR of the fMRI data. This is the matched filter principle. Although it is unknown, in practice, most research groups use FWHM ranges from 4mm to 10mm. A smoothing kernel of 6mm FWHM was used throughout our data.

fMRI Data Analysis

This is an area of active research. The field is moving ahead as more is learned about the physiological basis of fMRI, in particular the error structure of the stochastic process.

Panel Data. As fMRI data are typically panels, standard panel techniques can be used to draw inferences from the data. Each subject's data consists of a 3-dimensional time series. This space, which contains the brain, is divided into approximately 64×64×32 or ~130,000 voxels (volumetric pixels). The computational burden of data analysis is not trivial, therefore, and necessitates a number of simplifying assumptions in the data analysis.

Random effects analyses are usually preferred over fixed effects analysis in the case of fMRI data. This is because we are usually interested in drawing inference about the population, and it is therefore more appropriate to view the parameters as randomly distributed across the population.

Suppose there are N subjects, with each subject having T_i scans. The random effects models can be written as

$$y_i = X_i\beta + Z_i b_i + \varepsilon_i, \quad i = 1,\ldots,N$$
$$b_i \sim N(0,\Psi), \quad \varepsilon_i \sim N(0,\sigma^2\Omega)$$

where y_i is the dependent variable, X_i (size $T_i \times p$) and Z_i (size $T_i \times q$) are the fixed- and random-effects regressor matrices, and ε_i is the T_i-dimensional within subject error vector (this will be discussed later). The columns of Z_i will be a subset of the columns of X_i, as

there will in general be subject specific confounds or nuisance parameters (e.g., signal drift, motion parameters).

The random effects b_i can be thought of as individual deviations around the population β. Define b_i to have a mean of 0. Together with the assumption that the b_i's are drawn from a normal distribution, the random effects vector b_i is completely characterized by its variance-covariance matrix Ψ.

In a full mixed-effects analysis, the only restriction on Ψ is that it is symmetric and positive semi-definite. This, however, requires estimating more than 130,000 voxel-wise models, each with regressor matrices on the order of size $10{,}000 \times 10$ (assuming 20 subjects of 600 scans each with 10 regressors). To make this computation feasible, SPM makes the strong assumption that Ψ is an identity matrix. This has the effect of allowing the random effects model to be separated into a two-step procedure.

Step 1. Estimate each individual beta in $y_i = X_i\beta_i + \varepsilon_i$.

Step 2. Estimate second level analysis with $y = X\beta + \varepsilon$, where the vector y consists of the betas estimated from Step 1.

This two-step procedure in SPM is called a "summary statistic" procedure, and is essentially the result of two different assumptions. First, and fairly non-controversially, is that Ψ is a diagonal matrix. This results from the assumption that subjects' estimates are independent. Because each subject is in the scanner separately and can plausibly be assumed to have no influence on the other subjects in the study, this assumption seems fairly safe. The second assumption is that the variances of the subjects' estimates are identical. The effects of this assumption is not obvious and has been the topic of a number of papers (Friston et al. 1995c; Ashburner et al. 2001; Friston et al. 2005). Their findings suggest that the summary-statistic approach is generally adequate to justify the computational savings.

Despite the apparent simplicity of the random effects model, several issues must be dealt with. First, the shape of the hemodynamic response must be taken into account in the model. Second, the various sources of noise that enter into the error term (reviewed above), and especially autocorrelation of the error term, need to be adequately modeled to calculate meaningful *p*-values. Finally, because of the number of voxels in the brain (~130,000), a massive multiple comparisons problem is inherent in the analysis of fMRI data. Below we review some of the current approaches to these issues.

Hemodynamic Modeling. Because of the lagged nature of the BOLD response to neural activity, it is necessary to take into account the fact that the hemodynamic response is a lagged proxy of the underlying neuronal activity. One approach is to use a nonparametric approach and make no assumption about the underlying shape of the response. This can be done in what is called "event-related" fMRI studies in which a dummy variable is included at each TR of some event of interest (Glover 1999). One drawback of this approach is the proliferation of explanatory variables. More importantly from the point of view of statistical inference, by imposing minimal assumptions on the response, it is difficult to model parameterize this type of analysis in the cases when there is a strong a prior hypothesis.

The alternative, and the dominant method, is to assume some functional form for the response (HRF). SPM2 uses the sum of two gamma functions to capture the peak and the undershoot of the hemodynamic response (Friston et al. 1995c),

$$y(t) = A_1\left(\frac{x(t)-\delta_1}{\tau_1}\right)^2 \frac{e^{-\left(\frac{x(t)-\delta_1}{\tau_1}\right)^2}}{\tau_1} + A_2\left(\frac{x(t)-\delta_2}{\tau_2}\right)^2 \frac{e^{-\left(\frac{x(t)-\delta_2}{\tau_2}\right)^2}}{\tau_2} + C$$

where A_1 and A_1 are the magnitude of the peak and the undershoot, respectively; A_1 and A_1 model the width, peak height, and time to peak; A_1 and A_1 model the time to onset (Figure 13A). These parameters can, in addition, be modulated to vary the duration and/or intensity of the underlying neuronal response (Figure 13B).

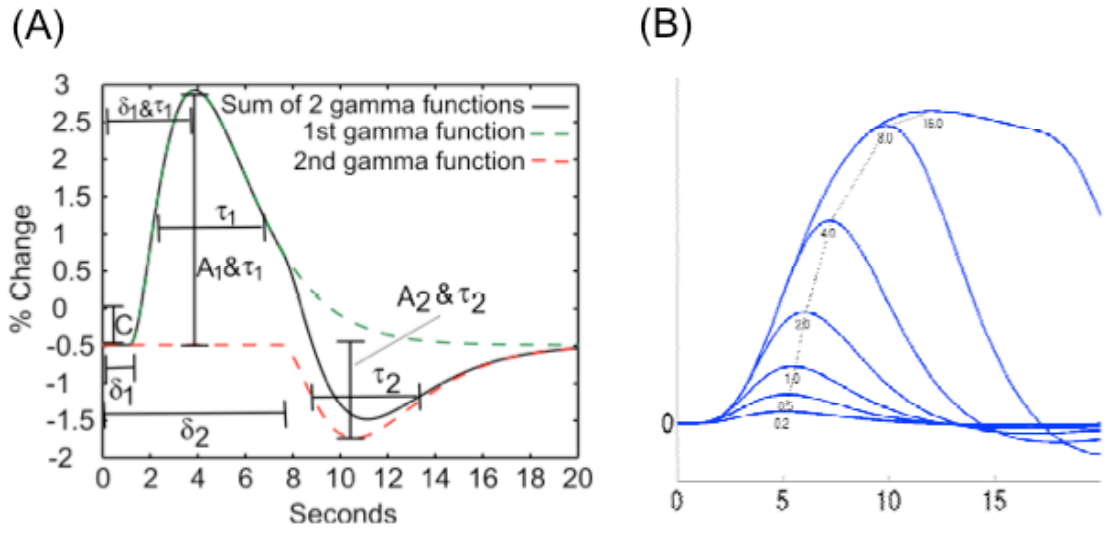


Figure 13: (A) Parameters for the sum of two gamma functions used to parameterize the hemodynamic response in SPM2. Adapted from Handwerker et al. (2004). (B) BOLD signal predicted from linear convolution by canonical impulse response of squarewave neural activity of increasing durations 200ms to 16s. Adapted from Henson (2001).

Autocorrelation. Autocorrelation in (with-in subject) fMRI data is captured in SPM2 by a combination of an AR(1) process and a highpass filter (Ashburner et al. 2001, Chapter 7). The AR(1) model captures short range correlations, with the longer-range correlations being captured by the high-pass filter. This requires the variance-covariance matrix for the AR(1) process to be estimated, which is potentially very computationally expensive if it is done voxel-wise. SPM2 makes the convenient assumption that the correlation structure is identical for all the voxels, but that the amplitude is different (Ashburner et al. 2001, Chapter 7). That is,

$$\Sigma_v = \begin{bmatrix} \sigma_v^2 & \rho\sigma^2 & \cdots & \rho^{T-1}\sigma_v^2 \\ \rho\sigma_v^2 & \sigma_v^2 & \cdots & \rho^{T-2}\sigma_v^2 \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1}\sigma_v^2 & \rho^{T-2}\sigma_v^2 & \cdots & \sigma_v^2 \end{bmatrix}$$

where ρ is the autocorrelation coefficient and σ_v^2 the variance of the voxel *v*. Making this assumption allows the autocorrelation estimate to be pooled over the voxels.

SPM2 estimates this AR(1) + white noise model using Restricted Maximum Likelihood (ReML), i.e.,

$$y_t = X_t\beta + \varepsilon_t + u_t$$
$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t$$

where $u_t \sim N(0,\sigma_u^2), v_t \sim N(0,\sigma_v^2)$, and ρ is the autocorrelation coefficient.

In sum, Figure 14 shows the steps that are used in the (within-subject) modeling of fMRI data. Panel A shows data (blue) and low-frequency drift (black) fitted by highpass filter (cutoff 168s); (B) Boxcar epoch model with (red) and without (black) convolution by a canonical HRF, after application of highpass filter; and (C) Residuals after fits of models with and without convolution. Note the systematic error for model without HRF convolution (black) at onset of each trial.

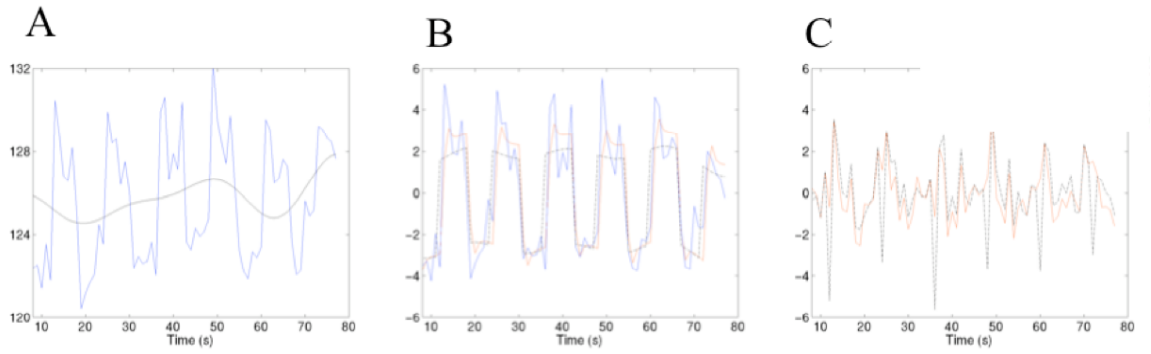


Figure 14: Highpass filtering, and HRF convolution. Adapted from Henson (2001).

CHAPTER 2

NEURAL SYSTEMS RESPONDING TO DEGREES OF UNCERTAINTY

In theories of choice under uncertainty used in social sciences and behavioral ecology, the only variables that should influence an uncertain choice are the judged probabilities of possible outcomes and the evaluation of those outcomes. Confidence in judged probability, however, can vary widely. In some choices, such as gambling on a roulette wheel, probability can be confidently judged from relative frequencies, event histories, or from an accepted theory. At the other extreme, such as the chance of a terrorist attack, probabilities are based on meager or conflicting evidence, where important information is clearly missing. The two types of uncertain events are often called risky and ambiguous, respectively. In subjective expected utility theory (SEU), the probabilities of outcomes should influence choices, but confidence about those probabilities should not. Many experiments show that many people are more willing to bet on risky outcomes than on ambiguous ones, holding judged probability of outcomes constant (Ellsberg 1961; Camerer and Weber 1992). This empirical aversion to ambiguity motivates a search for neural distinctions between risk and ambiguity, as in other studies on neuroeconomics, which explored the neural foundations of economic decision (Glimcher and Rustichini 2004; McClure et al. 2004; Camerer et al. 2005).

Ellsberg Paradox

The difference between risky and ambiguous uncertainty is illustrated by the Ellsberg paradox (32). Imagine one deck of 20 cards composed of 10 red and 10 blue cards (the risky deck). Another deck has 20 red or blue cards, but the composition of red and blue cards is completely unknown (the ambiguous deck). A bet on a color pays a fixed sum (e.g., $10) if a card with the chosen color is drawn, and zero otherwise (Fig. 1a).

In experiments with these choices, many people would rather bet on a red draw from the risky deck than a red draw from the ambiguous deck, and similarly for blue (Becker and Brownson 1964; MacCrimmon 1968). If betting preferences are determined only by probabilities, this pattern is a paradox. In theory, disliking the bet on a red draw from the ambiguous deck implies that its subjective probability is lower ($p_{amb}(red) < p_{risk}(red)$). The same aversion for the blue bets implies $p_{amb}(blue) < p_{risk}(blue)$. But these inequalities, and the fact that the probabilities of red and blue must add to one for each deck, implying that $1= p_{amb}(red) + p_{amb}(blue) < p_{risk}(red) + p_{risk}(blue) =1$, a contradiction. The paradox can be resolved by allowing choices to depend on subjective probabilities of events and on the ambiguity of those events. For example, if ambiguous probabilities are subadditive, then $1 - p_{amb}(red) - p_{amb}(blue)$ represents reserved belief and indexes the degree of aversion to ambiguity (Schmeidler 1989). Other models assume additive, but set-valued, probabilities, i.e., people believe that there is a range of possible probabilities, and ambiguity aversion is the result of people pessimistically assuming the worst probability (Gilboa and Schmeidler 1989). This model, and others, is silent about possible neural circuitry. Ambiguity aversion suggests that choices can depend on how much relevant information is missing, or how ignorant some people feel compared to others (Frisch and Baron 1988; Fox and Tversky 1995).

Ambiguity Aversion in Economics and Social Science.

Aversion to taking action in ambiguous situations has been studied in economics and politics (Mukherji and Tallon 2004), including macroeconomic policy-making (Sargent and Hansen), wage-setting and contracting (Mukerji 1998; Bewley 2002), strategic thinking (Camerer and Karjalainen 1994; Lo 1999), voting (40) (Ghirardato and Katz 2005), and financial investment (Dow and Werlang 1992; Epstein and Wang 1994). We illustrate with two examples from law and finance.

Law provides an interesting example that illustrates the psychology of ambiguity-aversion. In Scottish law there are three verdicts – guilty, not guilty, and not proven. The third is an unusual verdict in legal systems.[3] According to Peter Duff (1999), the difference between

"not guilty" and "not proven" is that the verdict of "not guilty" means that the accused definitely did not commit the crime. That is, "not guilty" is a positive declaration of innocence, whereas the verdict of "not proven" is thought to imply solely that the accused's guilt has not been conclusively demonstrated (p. 193). The "not proven" and "not guilty" verdicts have the same legal implication, because both prohibit retrial, even in the face of new evidence. "Not proven" verdicts are returned in about a third of jury trials, typically when the jury thinks the defendant is actually guilty but cannot legally convict because of a lack of corroborating evidence, which is required by Scottish law. For example, these verdicts are common in sexual assault trials where the only witness to the crime is the accusing victim and the jury believes the defendant is guilty, but cannot convict based on the weight of available evidence.

Turning to finance, "home bias" in investment is an important economic pattern which might be due to ambiguity-aversion. Home bias is the tendency for investors to invest in stocks that are, literally, closer to home. For example, investors in most countries tend to invest heavily in stocks from their home country and very little in stocks from foreign countries. In 1989 American, Japanese, and British investors held 94%, 98%, and 82% of their investments in home-country stocks (French and Poterba 1991), even though the latter two markets account for only a modest fraction of the world portfolio (

Figure 15). International home bias is shrinking however (Amadi 2004), as more investors buy global index funds and overseas stocks.

Unless investors have private information about their home stocks, home bias is a mistake because it leads to portfolios that are highly undiversified (especially for investors who do not live in the U.S. or Japan). Using 1989 data, French and Poterba estimate that given the apparent tradeoff between risk (stock return variation) and return (average percentage returns), the extra risk due to the reluctance to hold foreign stocks amounts to a sacrifice in annual percentage return of 1-2% per year (French and Poterba 1991). Assuming an average unbiased return of 7% (Siegel 1998), a typical historical estimate, a person with

home bias who invests a lump sum at age 25 will end up with only half as much money at age 65 as an investor who is unbiased and holds a worldwide index fund.

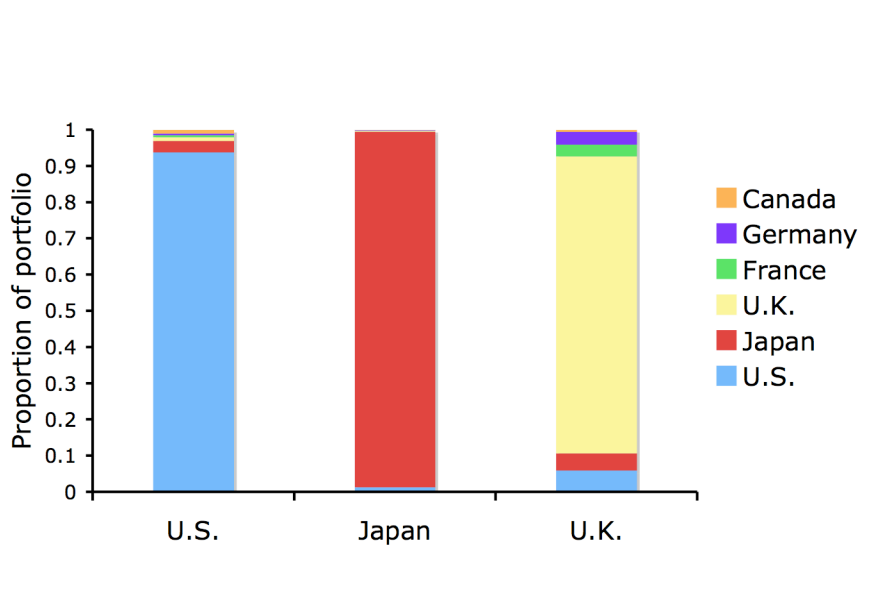


Figure 15: Portfolio weights for U.S., Japan, and U.K. in 1989. Data from French and Poterba (1991).

Interestingly, home bias exists at many levels besides the international one: Portfolio managers prefer to invest in companies with headquarters nearby, U.S. investors preferred their own regional "Baby Bell" companies after the breakup of AT&T, workers invest too heavily in the stock of companies they work for, and investors in many countries prefer to invest in nearby companies, or in those whose managers speak the same language that they do (Graham et al. 2006).

There are four commonly-evoked explanations of home bias: (i) higher transaction costs of buying foreign stocks; (ii) inside information about local stocks; (iii) optimism about relative returns of local stocks (Strong and Xu 2003); and (iv) aversion to ambiguity (which is usually called a taste for familiarity in finance research). A fifth explanation springs from the observation that patriotism across countries, measured by the World Values Survey, is correlated across countries with the extent of countrywide home bias (Morse and

Shive 2006). This is the national equivalent of preferring to bet on your home team in sports, but it does not seem to explain all the other levels of home bias.

The transaction cost and inside information explanations (i and ii) do not explain patterns like investment in the Baby Bell spin-off stocks. If investors are optimistic about local stocks and pessimistic about non-local stocks (explanation (iii)), they should short-sell the latter, but rarely do. (An investor who is ambiguity-averse toward non-local stocks will not want to buy them, and won't want to sell them short, either.) The familiarity explanation (iv), therefore, holds up rather well across all the levels of home bias that have been documented. This explanation is consistent with the idea that investors have a pure distaste for betting on either side of a proposition that they lack knowledge or familiarity about, which is very much like the knowledge treatment in our experiments.

Theory

Individuals, including economists, have a hard time understanding their own inconsistency with regard to ambiguity. Ellsberg wrote:

> *There are those who do not violate the axioms, or say they won't, even in these situations; such subjects tend to apply the axioms rather their intuition, and when in doubt, to apply some form of the Principle of Insufficient Reason. Some violate the axioms cheerfully, even with gusto; others sadly but persistently, having looked into their hearts, found conflicts with the axioms and decided, in Samuelson's phrase, to satisfy their preferences and let the axioms satisfy themselves. Still others tend, intuitively, to violate the axioms but feel guilty about it and go back into further analysis. (Ellsberg 1961)*

The standard way to think about choice under ambiguity is to assume that a person's value for a gamble is simply the average of the possible outcomes weighted by the probabilities that the outcome will occur. Specifically assume that there is a set of possible states of the world $e \in E$, each occurring with some probability (or with some probability distribution function if E is continuous), $p(e)$ and a set of actions $f \in L$, which yield some out come with utility $f(e)$ if e occurs. Expected utility theory states that the utility of some action $f \in L$ is equal to $\sum_{e \in E} p(e)f(e)$ (or the corresponding integral for continuous E). The existence of this sort of representation depends on the Independence Axiom stated below.

Axiom 3.1. *Independence Axiom* $\forall$ f, g, h $\in$ L, where L is the set of possible actions, f $\succ$ g $\Leftrightarrow$ $\alpha f + (1 - \alpha)h \succ \alpha g + (1 - \alpha)h \forall \alpha \in (0, 1)$.

Most alternatives to SEU focus on modifying this axiom to varying degrees. Several of these fall under the general rubric of Maxmin Expected Utility (MEU) (Gilboa and Schmeidler 1989). In these models, decision makers have a convex set of priors (a set of possible probability distributions over possible "states of nature") and act as if the worst case is realized.[4] The model uses a weakened version of the independence axiom, called "certainty-independence":

Axiom 3.2. Certainty-independence $\forall f, g \in L$ and $h \in L_c$, where L_c is the set of constant acts, and $\forall \alpha \in (0, 1), f \succ g$ iff $\alpha f + (1 - \alpha)h \succ \alpha g + (1 - \alpha)h$.

The standard independence axiom is stronger than certainty independence in that it allows h to be any act in L rather than restricting it to constant acts L_c, but both of these axioms explicitly deal with the reduction of compound lotteries, since this is at the center of how we represent ambiguity mathematically. Namely, there is some set of possible probability distributions over the states of nature, in the first stage of the compound lottery one of these distributions is chosen and in the second stage the actual state of nature is chosen. Under SEU this is equivalent to a one-stage lottery where the possible distributions are linearly

combined (in the discrete case the probability that some state $e \in E$ occurs is simply $\sum_P$ *Ł (P)P (s)* where *P* is a possible distribution and *Ł(P)* is the likelihood with which it occurs).

In addition, Gilboa and Schmeidler included an uncertainty aversion axiom, which states that decision makers prefer a mixture of acts with objective probabilities over the acts themselves.

$$\forall f, g \in L \text{ and } \alpha \in (0, 1), f \sim g \text{ implies } \alpha f + (1 - \alpha)g \succeq f.$$

Under standard conditions with the aforementioned changes, the preference relation is represented by a function *J(f)* up to a unique affine transformation,

$$J(f) = \min\left\{\int u \circ f dP \mid P \in C\right\}$$

where *f* is an act, *u* is a von-Neuman-Morgenstern utility function over outcomes, and *C* is a closed and convex set of finitely additive probability measures on the states of nature.

In these models, pessimism can be measured by the size of the set of priors. Larger sets of priors[5] will generally include worse possibilities and imply more pessimism. These models explain the Ellsberg paradox by allowing the probabilities assigned to each outcome (red or black), depend on the bet that is made, i.e., there is a set of possible values from *p(R)* ranging from p_* to p^*. When betting on red, the "worst" prior out of the decision maker's set is the one with the lowest odds on red, $p(R) = p_*$, so the expected utility of the red bet is low. When betting on black, the worst prior out of the set is the one with the lowest odds on black, $p(R) = p^*$ so $p(B) = 1 - p^*$, so the expected utility of the black bet is low. In this account, the sub-additivity of the revealed subjective probabilities in the paradox is due to the fact that <u>different</u> priors are used to evaluate the expected utility of each bet. This effect is like "Murphy's Law": if something can go wrong, it will.

Other models treat ambiguity as a two stage lottery where the actual probabilities are chosen in a first stage (Segal 1987). Here the explanation for ambiguity aversion comes from a violation of the reduction of compound lotteries. Non-expected utility is used to explain the Ellsberg paradox.

Finally, some models take the stance that ambiguity aversion is an overgeneralization of a rational aversion to asymmetry in information (Frisch and Baron 1988). These models argue that since many people confront incomplete information when they are facing a better-informed opponent, they treat the Ellsberg paradox as if there is asymmetric information (i.e., they act like they are playing a game against a malevolent experimenter who is trying to trick them). We examine this "informed opponent" hypothesis with the third treatment in our study, where they actually play against a better-informed opponent.

At some point in the future, economic theories might specify not only the link between unobservable factors and observed choices (such as beliefs and bet choices), but also a claim about neural circuitry that implements observed choices. Reviewing the many theoretical papers on ambiguity, we found one suggestion by Raiffa (1963) that echoes and anticipates the dual-systems models popular in behavioral economics today.

> *"But if certain uncertainties in the problem were in cloudy or fuzzy form, then very often there was a shifting of gears and no effort at all was made to think deliberately and reflectively about the problem. Systematic decomposition of the problem was shunned and an over-all 'seat of the pants' judgment was made which graphically reflected the temperament of the decision maker."*

Raiffa's suggestion seems to be that under ambiguity, deliberation and reflection (thought to be activities in prefrontal cortex) are limited, and a temperamental "seat of the pants" judgment takes over. Unfortunately, the seat of the pants is not a brain area, but we could

interpret him more broadly as suggesting a rapid emotional reaction to ambiguity. Thus, we translate Raiffa's observation, in neural terms, as implying a more rapid emotional reaction to ambiguous choices than to risky ones.

Experimental Design and Methods

Subjects

Sixteen Caltech undergraduate and graduate students were recruited from the Caltech SSEL database to participate in the study (13 males, 3 females). The mean (std. dev.) age was 23.5 years (6.2). Informed consent was obtained using a consent form approved by the Internal Review Board at Caltech. Subjects read written instructions before entering the scanner. After reading the instructions they completed a quiz to ensure comprehension of how their decisions affected their performance and earnings. They knew that at the end of the experiment, one trial from each of the three treatments would be chosen at random, and their choice on that trial would determine their pay. Their earnings were the total from the three randomly chosen choices, plus $5 fee for participating.

Behavioral Task

Stimuli were presented through MRI-compatible goggles (Resonance Technology). Choices were made using an MRI-compatible button box. For each choice, three options were given. Two of the options were bets on either side of a binary choice gamble that carried some uncertainty of paying either a positive sum or zero. The third option was the sure payoff that paid a certain positive amount of money. Subjects were allowed as much time as they desired in making their choice. Responses were made by pressing the button corresponding to the location of the options (left-middle-right) on the screen. The gambles were not played after each trial. This helps to eliminate any income effects as well as learning that might occur to reduce the degree of ambiguity over time.

Experimental Treatments

The fMRI study used three experimental treatments: Bets on card decks (based on the Ellsberg example above), bets on high-and low-knowledge world events, and bets against an informed or uninformed opponent (Figure 16).

The card deck treatment is a baseline pitting pure risk, where probabilities are known with certainty, against pure ambiguity. The knowledge treatment uses choices about events and facts which fall along a spectrum of uncertainty from risk to ambiguity. From the perspective of the MEU model, the high-knowledge questions correspond to smaller sets of priors than the low-knowledge question.

The informed opponent's treatment offers bets against another person, who is either better-informed or equally-informed about the contents of an ambiguous deck than the subject. In this condition the "opponent" draws a sample of cards from an ambiguous deck before making his bet. The subject has the option of making a simultaneous bet, but this will only count if the subject bets on the opposite color of the color the informed opponent chose. Since the opponent should always bet on the majority color from his sample, the subject's bet will only count when it is on the minority color from a random sample in the deck, i.e., when the subject is more likely to lose than win. This implies that the subject's expected value from winning is lower when the opponent is informed.

For example, assume that the subject has a prior that the deck will either contain 5 red cards and 15 blue cards or 15 red cards and 5 blue cards. Assume that the opponent draws one card before choosing how to bet, and the opponent always bets on the color that he drew. The subject, having no information about the deck, will bet randomly on red or blue. Suppose the subject chooses to bet on red (without loss of generality). When the deck has 5 red cards and 15 blue cards, the opponent will draw a blue with probability ¾. As a result, the subject will end up betting against the opponent's blue draw 75% of the time, and winning that bet only ¼ of the time. One-quarter of the time, the opponent draws a red card and bets red, in which case the subject's red bet does not count (and the subject earns the sure amount c). Therefore, the subject's expected value, conditional on the deck having 15 blue cards, is ¾ ¼x+ ¼c. The expected value conditional on the deck having 15 red cards is ¾c + ¼ ¾x. Since each of these conditional expected values is equally likely, the overall expected value is c/2 + 3/16 x. When the informed opponent does not have better information, then both subjects bet randomly and the subject's expected value is c/2 + x/4

(since half the time their colors match and they earn c, and half the time they bet and the subject wins that bet half the time). Comparing the two expected values, there is a small drop in expected value (x/4 - 3/16x = x/16) when betting against the informed opponent. This drop leads to a stricter constraint on when it is rational to take the gamble (x> 8/3c instead of x> 2c).

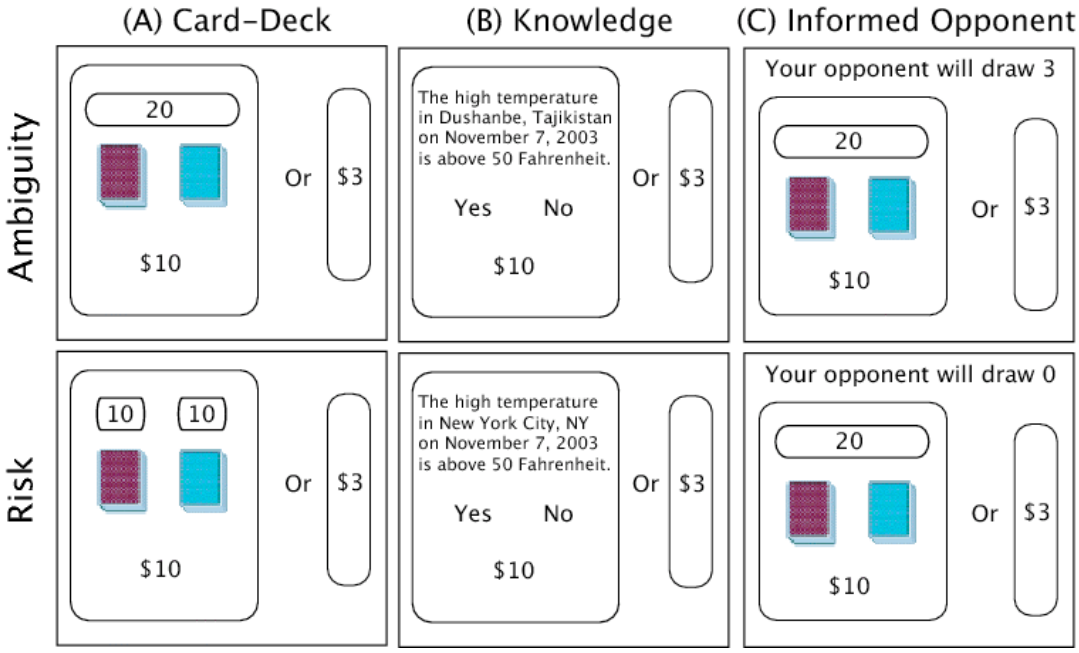


Figure 16: Sample screens from the experiment. (A) Bets on card decks (based on the Ellsberg example above); (B) bets on high- and low-knowledge world events; and (C) bets against an informed or uninformed opponent.

Notice that all three treatments have one condition where the subjects are missing some relevant information relative to the other. We call all these treatments the ambiguous conditions and the other treatments the risky conditions. Subjects made 24 choices in each treatment between certain amounts of money and bets on events. The amount of the certain payoff and the bet varied across trials.

In these experiments we allow 6-8 seconds between trials. This break was necessary to allow the activations caused by the previous trial to dissipate. We also randomize the length

of these inter-trial intervals, because using a fixed interval can create anticipation effects in the seconds just before the new trial is presented; having a random inter-trial interval length diminishes these effects.

fMRI Data Collection

Imaging was performed using a 3 Tesla Siemens Trio scanner at the Broad Imaging Center at Caltech. A set of high-resolution (0.5mm x 0.5mm x 1.0mm) T1-weighted anatomical images was first acquired to enable localization of functional images. Whole-brain functional images were acquired in 32-34 axial slices (64 x 64 voxels; in plane resolution 3 mm x 3 mm x 3.5 mm slices, number of slices varies according to variation in the subjects' brain sizes) at a TR of 2000 msec, TE of 30 msec.

The scan sequences were axial slices approximately parallel to the AC-PC axis. Scan sequences were not optimized for the OFC, therefore susceptibility artifacts affected adversely the image quality of the medial OFC. The lateral OFC activation found in the experiment, however, showed no such signal dropout upon inspection.

Prior to analysis, the images were corrected for slice time artifacts, realigned, coregistered to the subject's T1 image, normalized to Montreal Neurological Institute coordinates (resampled 4mm x 4mm x 4mm), and smoothed with an 8mm full-width-at-half-maximum Gaussian kernel using SPM2 (for details see Chapter 1). At the start of each functional scanning run, the screen remained black for 4s to allow time for magnetization to reach steady state. The associated first two images were discarded from the analysis. Thirteen out of 16 subjects completed all tasks within three 15-minute scanning runs. However, the duration of the experiment was variable because choices were self-paced. Three subjects did not have time to complete the Informed Opponent treatment.

fMRI Data Analysis

For each treatment, we estimated a general linear model (GLM) using standard regression techniques (for details, see Chapter 1). Two primary regressors were used in the GLM, one for the ambiguity trials and one for the risky trials. The regressors were constructed in the

following way: first we created a boxcar regressor (dummy variable) that was 1 during the risky (ambiguous) trial and 0 elsewhere. These regressors were then convolved with the hemodynamic response function. Maximum likelihood estimation was used to fit:

$$B_t = \beta_{amb} A_t + \beta_{risk} R_t + \beta_0$$

where B_t is the time series for some voxel in the brain[6]. Voxels with regression coefficients significantly different from 0 can be said to covary with either risk or ambiguity. We are mostly interested in which voxels are differentially activated by ambiguity with respect to risk, i.e. voxels where β_{amb} is significantly different from β_{risk}.

As the experiment was self-paced, the length of the trials varied (see behavioral data in results section for details). To find areas differentially activated by ambiguity and risk across all three treatments, a one-way ANOVA of the three conditions was performed, correcting for non-sphericity, and excluding areas that are significantly different between the three treatments.

The regressors were anchored to stimulus presentation, i.e., the original dummy variable turns "on" when the stimulus is presented, as opposed to when a decision is made, based on the hypothesis that the reaction to uncertainty would occur before the decision. To check the robustness of our results to this assumption, we also analyzed the time courses of activation anchored at both the stimulus onset and the decision time.

Some models in the analysis used the expected value of subjects' choices. For the ambiguous gambles and for gambles where no objective probabilities were given (e.g., ambiguous gambles in all three treatments, and risky gambles in Knowledge treatment), we calculated the expected value by assuming that the probabilities were equiprobable. Furthermore, for the Informed Opponents hypothesis, we invoked the additional assumption that subjects have a prior distribution over the cards (we used both a uniform and binomial prior; details available in endnote 7).

Behavioral Results

Summary of Behavioral Results

Table 1 shows mean response times across conditions and treatments. Because response times are skewed and bounded away from zero, standard deviations are large and misleading about distributional variance. Figure 17, therefore, displays boxplots of the log response times of choices by type of choice, condition, and treatment. These boxplots show that there are no significant differences in the central tendency of response times between any of these variables.

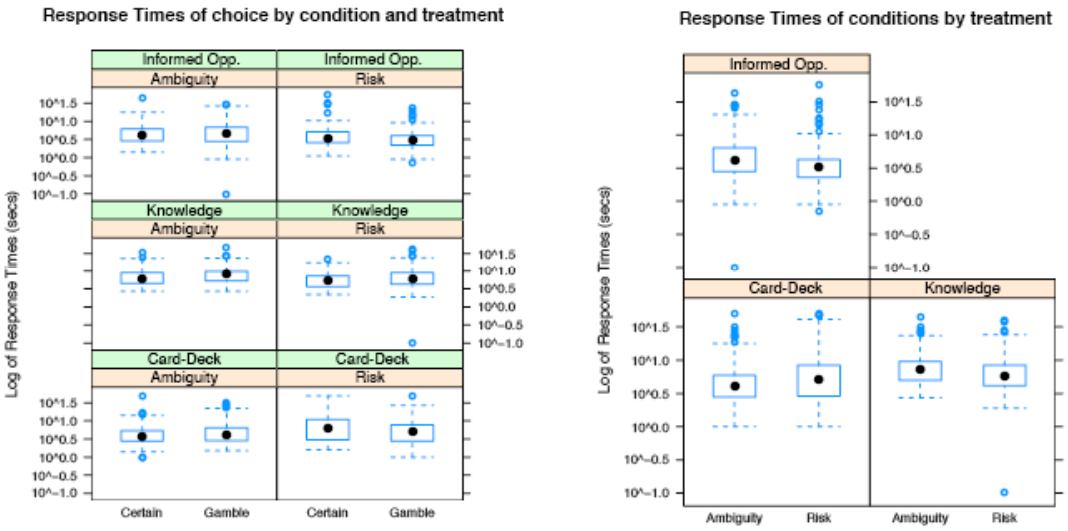


Figure 17: Boxplots of log response times (secs) of choice by condition and treatment. Response times are logged due to skewness of the distribution. Boxplots show that there are no significant differences between choices and between conditions across the treatments.

	Ambiguity		Ambiguity Total	Risk		Risk Total	Grand Total
Treatment	Certain	Gamble		Certain	Gamble		
Card-Deck	4.88	5.67	5.30	8.45	6.30	7.13	6.22
Knowledge	7.53	8.29	8.03	6.03	7.32	6.98	7.51
Informed Opp.	5.39	5.93	5.69	5.40	3.54	3.96	4.83
Grand Total	5.83	6.80	6.39	7.00	5.80	6.16	6.27

Table 1: Summary of mean response times (secs) of choices across treatments.

Summary of the proportion of choices made by subjects are presented in Figure 18 and Table 2. Roughly speaking, a greater number of certain choice indicates greater ambiguity/risk aversion. A more rigorous demonstration of ambiguity/risk aversion is presented later.

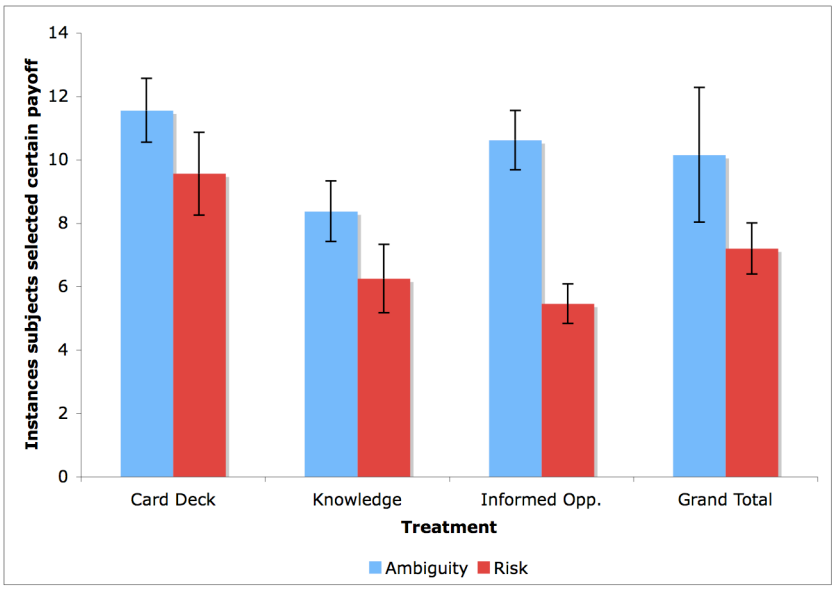


Figure 18: Choices of subjects across treatment and condition. Average of number and standard error bars of certain payoff choice by subjects across condition and treatment (out of 24 total choices per condition per treatment). A higher frequency of choosing the certain payoff in the ambiguity treatment (compared to the frequency in the risk treatment) is an indicator of ambiguity aversion.

Condition/Treatment	*Card-Deck*	*Knowledge*	*Informed Opp.*	*Grand Total*
Ambiguity	11.56	8.37	10.61	10.15
Risk	9.56	6.25	5.46	7.20
Ambiguity – Risk	2.00	2.12	5.15	2.95

Table 2: Summary of subject choices across treatment and condition. Average number of certain payoff choices by subjects across conditions and treatments (out of 24 total choices per condition per treatment).

Risk and Ambiguity Attitudes

We estimate subjects' risk and ambiguity attitudes via a stochastic choice model. The subjects' utility functions for money are assumed to follow a power function $u(x, \rho) = x^{\rho}$.

The power function is conveniently characterized by one parameter and widely employed in empirical estimation of choice under uncertainty. Subjects are assumed to weight probabilities according to the function $\pi(p, \gamma) = p^{\gamma}$. The ρ parameter is interpreted as the risk aversion coefficient, i.e., the curvature of the utility function. The γ parameter is interpreted as the ambiguity aversion coefficient, i.e., how much do people over(under)-weight probabilities because they are not confident in their judgments. If subjects over-weight ambiguous probabilities ($\gamma < 1$), we characterize them as ambiguity-preferring. If they under-weight ambiguous probabilities ($\gamma > 1$, as in the non-additive prior view), we characterize them as ambiguity-averse. If subjects weight probabilities linearly ($\gamma = 1$), we characterize them as ambiguity-neutral. We assume subjects combine these weighted-probabilities and utilities linearly, so that their weighted subjective expected utility is $U(p, x; \gamma, \rho) = \pi(p, \gamma)u(x, \rho)$.

The tasks are all binary choices in which subjects either choose a gamble to win x (with probability p) or 0, or a certain payoff c. For the risky deck, the ratios of the cards are the probabilities. For the ambiguous decks and all knowledge questions, we assumed $p = 1/2$. If subjective p is different than 1/2 (e.g., because a subject happens to know a lot about fall temperatures in New York), then subjective probabilities are not held constant across the knowledge trials. This possibility biases our analysis against finding common regions of activation across treatments, so would imply that the results described in the text are conservative about the true extent and commonality of ambiguity and risk-specific regions. We constrain $\gamma = 1$ in all risk conditions and estimate γ from behavioral data in the ambiguity conditions.

The probability that the subject chooses the gamble rather than the sure amount c is given by the logit or softmax formula,

$$P(p,x;\gamma,\rho) = \frac{1}{\left(1 + \exp\{-\lambda(U(p,x;\gamma,\rho) - u(c,\rho)\right)}.$$

The parameter λ is the sensitivity of choice probability to the utility difference (the degree of inflection), or the amount of "randomness" in the subject's choices (λ =0 means choices are random; as λ increases the function is more steeply inflected at zero).

Denote the choice of the subject in trial *i* by y_i, where y_i =1 if subject chooses the gamble, and 0 if the subject chooses the certain payoff. We fit the data using maximum likelihood, with the log likelihood function

$$\sum_{i=1}^{N} y_i \log\big(P(p,x,c;\gamma,\rho,\lambda)\big) + (1-y_i)\log\big(1-P(p,x,c;\gamma,\rho,\lambda)\big).$$

Because this is a nonlinear optimization problem, numerical methods must be used. We used the Nelder-Mead simplex algorithm (Nelder and Mead 1965) implemented in Mathematica v5.1, with 10 random starting positions. The iteration with the highest likelihood value was chosen. This procedure is quite robust as the estimates are in general insensitive to the starting values. Table 3 shows the estimates from the fMRI subjects.

Subject	**Card-Deck**			**Knowledge**		
	γ	ρ	λ	γ	ρ	λ
AJB	0.98 (0.10)	1.00 (0.09)	1.39 (0.63)	1.11 (0.15)	1.13 (0.12)	0.71 (0.41)
APS	0.80 (0.04)	0.90 (0.04)	4.32 (1.92)	1.66 (0.57)	1.44 (0.48)	0.10 (0.18)
BIC	0.81 (0.11)	0.73 (0.08)	1.81 (0.73)	1.05 (0.22)	0.63 (0.11)	1.37 (0.60)
BSU	1.20 (0.10)	0.73 (0.04)	9.98 (4.82)	1.09 (0.19)	1.07 (0.15)	0.51 (0.34)
CSJ	0.93 (0.07)	0.87 (0.04)	3.85 (1.50)	1.30 (0.46)	1.50 (0.71)	0.10 (0.26)
DAS	0.84 (0.19)	0.85 (0.18)	0.85 (0.81)	1.70 (1.13)	1.29 (0.84)	0.10 (0.33)
EJH	0.82 (0.12)	0.83 (0.10)	1.39 (0.67)	1.09 (0.17)	1.05 (0.13)	0.64 (0.39)
HCH	0.55 (0.08)	0.52 (0.06)	6.65 (2.64)	1.52 (0.12)	1.25 (0.10)	1.43 (0.89)
KED	1.03 (0.05)	1.05 (0.04)	3.98 (1.67)	1.19 (0.18)	1.27 (0.17)	0.48 (0.33)
LTL	0.98 (0.15)	1.02 (0.14)	0.80 (0.55)	1.07 (0.09)	0.89 (0.05)	2.13 (0.78)

MK	1.12 (0.16)	1.01 (0.13)	0.85 (0.52)	1.30 (0.45)	1.37 (0.55)	0.10 (0.21)
PRV	1.22 (0.14)	1.13 (0.12)	1.22 (0.69)	1.33 (0.23)	1.47 (0.22)	0.33 (0.26)
SWT	0.81 (0.09)	0.58 (0.05)	5.01 (1.56)	1.10 (0.06)	0.70 (0.03)	10.16 (3.77)
TEJ	0.84 (0.06)	0.79 (0.03)	5.35 (1.83)	1.49 (0.26)	1.36 (0.22)	0.26 (0.21)
VS	1.19 (0.19)	1.22 (0.18)	0.63 (0.51)	1.42 (0.27)	1.50 (0.26)	0.15 (0.14)
WL	0.85 (0.13)	0.83 (0.10)	1.11 (0.59)	1.88 (0.41)	1.38 (0.27)	0.18 (0.18)
Mean est. Mean s.e.	0.94 (0.11)	0.88 (0.09)	3.07 (1.35)	1.33 (0.31)	1.21 (0.28)	1.17 (0.58)

Table 3: Ambiguity aversion estimates. Separate estimates (standard errors) of the ambiguity aversion coefficient γ, risk aversion coefficient ρ, and inflection parameter λ for subjects in the Card-Deck and Knowledge Conditions.

fMRI Results

Ambiguity and Risk Regions

Areas that were more active during the ambiguous condition relative to the risk condition are listed in Table 4. These included OFC and amygdala (Figure 19a, b). Critchley et al. (2001) found OFC activation as subjects anticipated information about a financial gain or loss. Our study shows activity in this area even though there is no feedback during the experiment. More subtly the OFC appears to be involved in integrating emotional and cognitive information. OFC lesion patients often behave inappropriately in social situation despite knowledge of what proper behavior entails (Berthoz 2002), so the OFC may be active generally in emotional integration over a wide spectrum of situations.

The amygdala has been specifically implicated in processing information related to fear, e.g., recognizing frightened faces (Bechara, 2003, Adolphs, 2002, Critchley, 2000}. We hypothesized that this area would be important for processing uncertain events since risk and ambiguity aversion could be interpreted as fear of the unknown. The amygdala is also involved in emotional learning and conditioning, both of which should be relevant in dealing with ambiguous situations (Phelps et al. 2004). In addition, the role of the amygdala in reacting to missing information is consistent with evidence of its involvement

in interpersonal evaluations with missing social information, where familiarity modulates amygdala response (Phelps et al. 2000). Showing unfamiliar black faces to white subjects elicits amygdala activation which correlates with the strength of implicit associations between black names and negatively-valenced words (Phelps et al. 2000). This correlation disappears when the black faces are familiar (e.g., Bill Cosby), which suggests the amygdala may be partly reacting to ambiguity of the social evaluation of faces; activity dissipates when faces are familiar and more like risky gambles which have less missing social information. Similarity of activity in all three fMRI treatments also suggests that aversion to betting on ambiguous events may be an overgeneralization of a rational aversion to betting against other agents who are better informed.

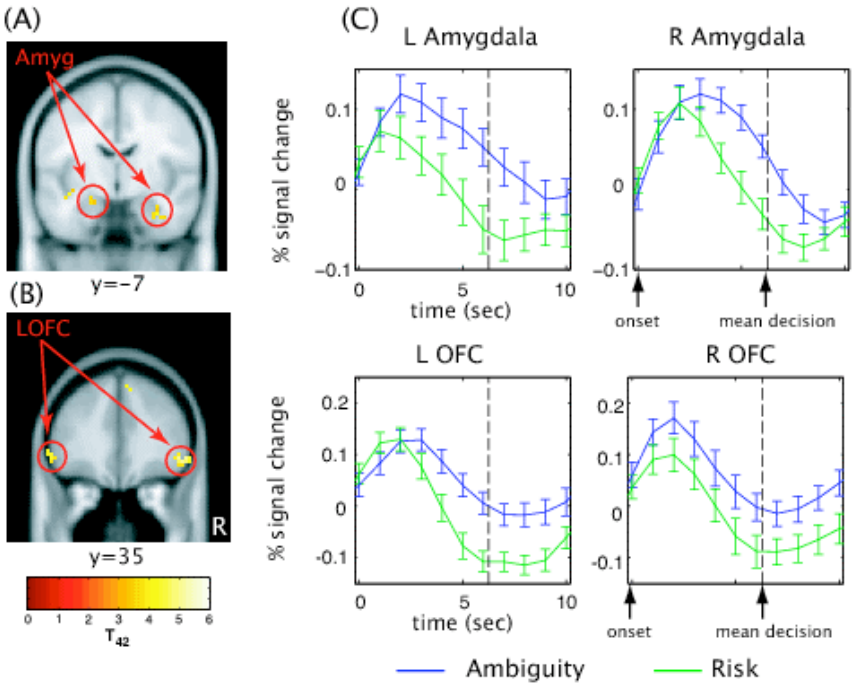


Figure 19: Regions showing greater activation to ambiguity than risk: Random effects analysis of all three treatments revealed regions that are differentially activated in decision making under ambiguity relative to risk (at p . 0.001, uncorrected). These regions include (A) left amygdala, and right amygdala/parahippocampal gyrus, (B) bilateral OFC, and bilateral inferior parietal lobule (Table S2). (C) Mean time courses of amgydala and OFC (time synched to trial onset, dashed vertical lines are mean decision times; error bars are SEM; n=16). Time courses are plotted using the most signi.cant voxel in each cluster.

Areas activated during the risk condition relative to ambiguity are listed in Table 5. These include the dorsal striatum (caudate nucleus) (Figure 20a), an area that has been implicated in reward prediction (Schultz 2000; Knutson et al. 2001; O'Doherty et al. 2004). One earlier study also found differential activation in the caudate during risk relative to ambiguity using PET (Rustichini et al. 2005).[8]

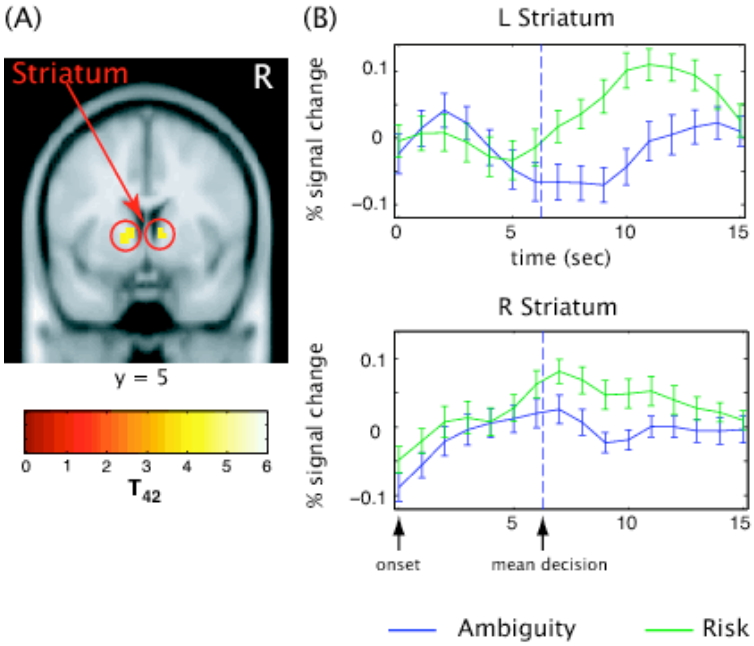


Figure 20: Regions showing greater activation under risk than ambiguity: Random effects analysis of all three treatments revealed brain regions that are differentially activated in decision making under risk. These regions include (A) dorsal striatum, and also precuneus and premotor cortex (Table S3). (B) Mean time courses for risk areas (time synched to trial onset, dashed vertical lines are mean decision times; error bars are SEM; n=16).

Time courses also showed different patterns of activation in the ambiguity > risk and risk > ambiguity regions indicating two distinct systems at work. Whereas the amygdala and OFC reacted rapidly at the onset of the trial (Figure 19b), the dorsal striatum activity built more slowly and peaks after the decision time (Figure 20b). Furthermore, these activations are present in all three experimental treatments (see Figure 21).

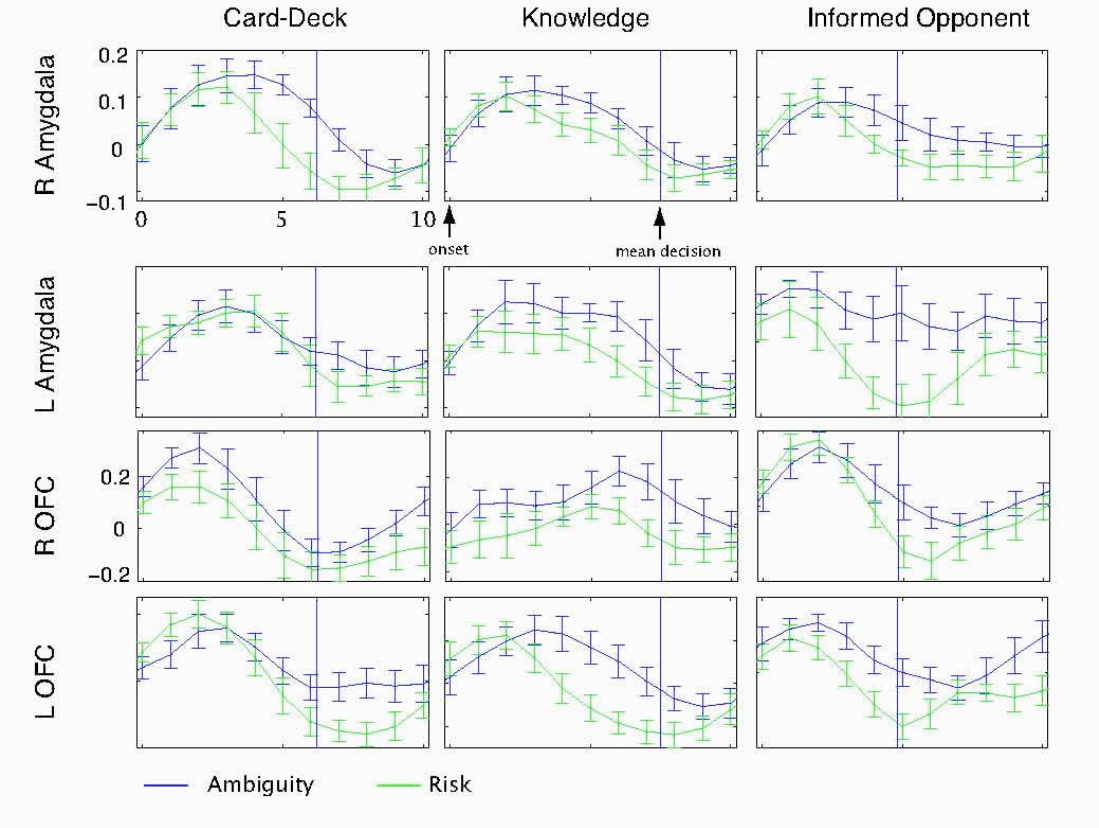


Figure 21: Time courses of percentage signal change in brain regions that are differentially activated in decision making under ambiguity in Card-Deck, Knowledge and Informed Opponent conditions. Note that the qualitative aspects of the activation differences between ambiguity and risk are preserved between the pooled time courses in Figure 2.

cluster			voxel							Regions	
p_{cor}[1]	k_E[2]	p_{unc}[3]	p_{FWE}[4]	p_{FDR}[5]	T[6]	Z[7]	X[8]	Y	Z	L/R[9]	Region
0.01	82	0.001	0.011	0.007	5.96	5.04	51	33	-6	R	Lateral Orbitofrontal Cortex

[1] Corrected (family-wise) cluster-level *p*-value.

[2] Cluster size (voxels).

[3] Uncorrected cluster-level *p*-value.

[4] Corrected (family-wise) voxel-level *p*-value.

[5] Corrected (false-discovery rate) voxel-level *p*-value.

[6] T-statistic of voxel.

[7] Z-score of voxel.

[8] (X, Y, Z) are the MNI coordinate of voxel location (mm).

[9] Laterality (L = left, R = right).

			0.897	0.017	3.92	3.6	54	18	-21		
0	109	0	0.052	0.007	5.38	4.67	-54	-60	42	L	Inferior Parietal Lobule
			0.1	0.007	5.13	4.5	-45	-54	33		
0	112	0	0.06	0.007	5.33	4.63	-9	48	39	L	Dorsomedial Prefrontal Cortex
			0.306	0.008	4.66	4.16	-12	63	21		
0	119	0	0.072	0.007	5.26	4.59	54	-54	36	R	Supramarginal Gyrus
			0.599	0.01	4.3	3.89	54	-63	30		
0	226	0	0.162	0.007	4.94	4.36	18	54	18	R	Dorsomedial Prefrontal Cortex
			0.229	0.008	4.79	4.26	12	54	30		
			0.379	0.009	4.56	4.09	12	27	57		
0.06	52	0.007	0.201	0.008	4.85	4.3	36	18	42	R	Middle Frontal Gyrus
			0.884	0.016	3.94	3.62	42	9	45		
0	154	0	0.22	0.008	4.81	4.27	60	-36	-3	R	Middle Temporal Gyrus
			0.485	0.009	4.43	3.99	63	-27	-6		
			0.626	0.01	4.27	3.87	51	-24	-9		
0.44	21	0.066	0.302	0.008	4.67	4.17	-39	-9	-15	L	Sub-Gyral
0.13	40	0.015	0.331	0.009	4.63	4.14	39	6	-27	R	Frontoinsular Cortex
			0.951	0.019	3.8	3.5	42	15	-24		
0.41	22	0.061	0.547	0.01	4.36	3.94	54	27	6		Lateral Orbitofrontal Cortex
0.26	29	0.034	0.584	0.01	4.32	3.91	-54	36	-6	L	Lateral Orbitofrontal Cortex
0.74	12	0.154	0.75	0.013	4.13	3.76	-15	-15	-15	L	Amygdala/Parahippocampal Gyrus
			0.993	0.026	3.57	3.32	-21	-6	-18	L	Amygdala
0.41	22	0.061	0.825	0.014	4.03	3.69	33	-6	-27	R	Amygdala/Parahippocampal Gyrus

Table 4: *Ambiguity* > *Risk* regions. Local maxima of clusters, $p < 0.001$ uncorrected, clusters with $k < 10$ voxels not shown (All local maxima uncorrected p-values are significant to three significant figures, and are omitted from the table.)

cluster			voxel							Regions	
p_{cor}	k_E	p_{unc}	p_{FWE}	p_{FDR}	T	Z	X	Y	Z	L/R	Region
0.06	52	0.007	0.063	0.012	5.31	4.62	0	-6	6	M	Caudate
			0.993	0.033	3.57	3.32	9	6	6	R	
			0.952	0.023	3.79	3.5	-12	6	0	L	
0	641	0	0.07	0.012	5.27	4.59	12	-60	-3	R	Culmen
			0.119	0.012	5.07	4.45	9	-78	3	R	Lingual Gyrus
			0.162	0.012	4.94	4.36	-12	-75	15	L	Cuneus
0.01	81	0.001	0.295	0.012	4.68	4.18	-15	-72	51	L	Precuneus
0.26	29	0.034	0.338	0.012	4.62	4.13	-3	9	45	L	Precentral Gyrus
0.12	41	0.014	0.569	0.012	4.33	3.92	12	-75	51	R	Precuneus
			0.906	0.02	3.9	3.58	21	-84	39		
0.74	12	0.154	0.923	0.021	3.87	3.56	-42	-75	30	L	Angular Gyrus

Table 5: *Risk*>*Ambiguity* regions. Local maxima of clusters, $p < 0.001$ uncorrected, clusters with $k < 10$ voxels not shown (All local maxima uncorrected p-values are significant to three significant figures, and are omitted from the table.)

Note that it is rational in the informed opponent's treatment to discount some of the payoff in the gamble, as the gamble only wins if the better-informed opponent chosses the wrong color. The informed opponent hypothesis is that bets on ambiguous card decks and low-knowledge events, while normatively different than bets against the informed opponent, are generated by similar neural circuitry. The time courses in the amygdala, OFC, and striatum are similar across all three treatments, consistent with this hypothesis.

One simplified way of interpreting this data is to hypothesize that there are two interacting systems: (1) a vigilance system in the amygdala and OFC that responds more rapidly to the stimuli, and (2) a reward-anticipation system in the striatum that is further downstream. The overall activity differences in the contrasts indicate that system (2) is more active during risky decisions, which makes sense since, in these situations, subjects have the information necessary for accurate reward prediction. Conversely, during ambiguous conditions, the first rapid system (1) appears to be more active, indicating that it may be reacting to the level of information available (less information during the ambiguous situations leads to greater vigilance in the form of higher activation levels in the amygdala and OFC). Both systems are active to varying degrees during risky and ambiguous trials, the difference is one of degree—as the level of information available to the decision maker rises, activity in system (1) declines relative to more ambiguous situations; the converse is true of system (2).

Expected Reward Regions

To identify regions that were sensitive to expected dollar value of rewards, conditional on the subject's choice, we used a model where the event of interest was synched to the decision epoch.

Two different main regressors were used for the ambiguity and risk conditions, respectively. Each regressor was associated with an interaction term defined by the expected dollar value of the actual choice (either the gamble or the certain payoff). Notice that sometimes the choice was the certain payoff and sometimes it was the gamble, so this

interaction should detect regions that are sensitive to expected value of the actual choice, regardless of whether it was a certain or uncertain amount.

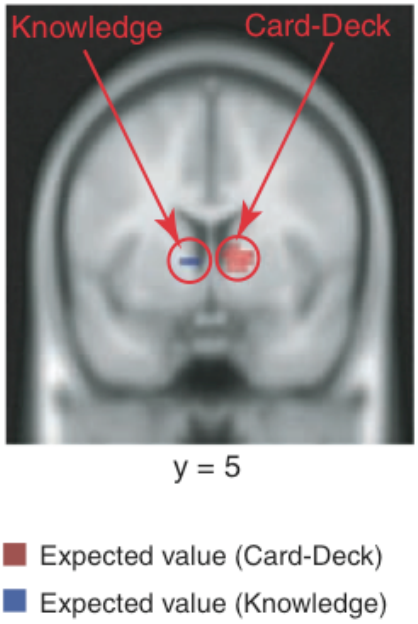


Figure 22: Regions of the dorsal striatum significantly correlated with expected values of subjects' choices in risk condition of Card-Deck treatment (red) and both risk and ambiguity conditions of Knowledge treatment (blue) ($p < 0.005$, uncorrected; cluster size $k \geq 10$ voxels).

Results from this analysis are presented in Table 6 and Table 7. In the Card-Deck treatment, activity in the right dorsal striatum was correlated with the expected reward in the risk condition

Figure 22. Activity in left dorsal striatum was correlated with the expected reward in both conditions of the Knowledge treatment. No consistent activations were found in the Informed Opponents treatment (which is likely due to the sensitivity of the calculation of expected value in this treatment, see Endnote 7).

Importantly, the dorsal striatum regions found in this analysis also overlapped with the *risk* > *ambiguity* regions found in the main analysis. In fact, it is the unique cluster that exhibits

this property. In addition, once we take the expected value of the choices into the model, the dorsal striatum is no longer significantly activated by risk. These results strongly suggest that the difference in the time course is due to expected reward differences.

Finally, it is interesting to note that the lateralization corresponds to the semantic/mathematical lateralization. It is less clear, however, how much of this dissociation is due to the ambiguity/risk distinction, or the lack of information about the expected reward in the ambiguity condition. Further research is needed to establish this relationship.

cluster			voxel								Regions	
p_{cor}	k_E	p_{unc}	p_{FWE}	p_{FDR}	T	Z	p_{unc}	X	Y	Z	L/R	Region
0.002	120	0	0.089	0.089	7.11	4.64	0	9	24	54	R	Superior Frontal Gyrus
			0.994	0.397	4.85	3.71	0	12	39	51		
			1	0.397	4.47	3.51	0	9	12	54		
0.429	36	0.01	0.32	0.16	6.36	4.36	0	60	-33	3	R	Middle Temporal Gyrus
			1	0.397	3.31	2.82	0.002	57	-42	3		
0.844	23	0.032	0.781	0.314	5.76	4.12	0	-51	-72	30	L	Angular Gyrus
			1	0.397	3.99	3.24	0.001	-42	-78	33		
1	10	0.139	0.991	0.397	4.92	3.74	0	-66	-30	-9	L	Middle Temporal Gyrus
0.487	34	0.012	0.996	0.397	4.81	3.68	0	-9	-18	18		
			1	0.397	4.16	3.34	0	-9	-3	18		
0.403	37	0.009	1	0.397	4.55	3.55	0	15	6	3	R	Caudate
0.041	70	0.001	1	0.397	4.37	3.46	0	48	24	-15	R	Inferior Frontal Gyrus
			1	0.397	3.67	3.05	0.001	54	39	-3	R	Brodman Area 47
			1	0.397	3.54	2.97	0.001	39	24	-15		
0.783	25	0.026	1	0.397	4.21	3.37	0	-9	45	9	L	Anterior Cingular Gyrus
0.549	32	0.014	1	0.397	3.99	3.24	0.001	-3	57	33	L	Brodman Area 9
			1	0.397	3.87	3.17	0.001	-9	39	30		
			1	0.397	3.46	2.92	0.002	-6	48	27		
0.999	11	0.122	1	0.397	3.95	3.22	0.001	3	-42	-6	M	Culmen
0.897	21	0.039	1	0.397	3.85	3.16	0.001	60	-57	30	R	Supramarginal Gyrus
			1	0.397	3.73	3.09	0.001	60	-51	39		
0.999	11	0.122	1	0.397	3.84	3.16	0.001	-6	27	42	L	Brodman Area 6
1	10	0.139	1	0.397	3.74	3.1	0.001	-45	42	6	L	Inferior Frontal Gyrus
1	10	0.139	1	0.397	3.43	2.9	0.002	21	48	27	R	Superior Frontal Gyrus
			1	0.397	3.38	2.87	0.002	18	57	33		

Table 6: Regions positively correlated with expected value of decisions in risk condition of Card-Deck treatment. Local maxima of clusters, $p < 0.005$ uncorrected, clusters with $k < 10$ voxels not shown (All local maxima uncorrected).

cluster			Voxel								region	
p_{cor}	k_E	p_{unc}	p_{FWE}	p_{FDR}	T	Z	p_{unc}	X	Y	Z	L/R	Region

0.782	41	0.06	0.417	0.741	5.88	4.17	0	33	-87	15	R	Middle Occipital Gyrus
			1	0.93	3.53	2.96	0.002	39	-84	9		
0.995	17	0.21	0.998	0.93	4.08	3.3	0	-9	12	3	L	Caudate
			1	0.93	3.24	2.78	0.003	-6	3	6		
1	10	0.334	1	0.93	3.69	3.06	0.001	-24	-63	51	L	Superior Parietal Lobule

Table 7: Regions positively correlated with expected value of decisions in Knowledge treatment. Local maxima of clusters, $p < 0.005$ uncorrected, clusters with $k<10$ voxels not shown (All local maxima uncorrected).

Decision-Synched Models

A model with regressors anchored to the decision epoch was also estimated for all subjects. This provides a robustness check on our assumption that the hemodynamic responses in our regions were synched to the onset of the stimulus, rather than that of the decision. Results from this model showed similar activation in the dorsal striatum for the risk > ambiguity contrast, but did not show differences in the amygdala or lateral OFC. The hemodynamic responses in Figure 19 and Figure 20 clearly show why: the activations in the OFC and amygdala occur at the beginning of the trial, and peaks before the decision epoch (Figure 19); the striatal activity occurs somewhere between the onset and the decision, which allows it to be captured by the decision-synched model (Figure 20).

Choice-Dependent Regions

Table 8 and Table 9 present all regions differentially activated under gamble and certain payoff choices ($p < 0.001$ uncorrected, $k \geq 10$ voxels). In addition to contralateral visual and motor activations corresponding to the visual inputs and motor responses required to make the choices, there were significant bilateral insula and left ventral striatum activation in the gamble>certain contrast.

This is consistent with previous findings of insular activity in decision making under risk (Paulus et al. 2003). These regions, however, did not exhibit significant interaction with the ambiguity/risk distinction. Together with the fact that the risk/ambiguity regions did not exhibit significant differences across choices, this suggests that the risk/ambiguity regions are indeed responding to the ambiguity/risk trial dimension rather than to gamble/certain choices.

Notably, there do not appear to be any interactions between these decision-conditions and the ambiguity vs. risk conditions. This implies that the differences we see in the Ambiguity/Risk comparisons are purely the reaction to the level of uncertainty the subject is exposed to. While these reactions are almost certainly, then, an input into the actual decision, the areas we see appear to be involved in evaluation of the situation rather than determining choice.

cluster			Voxel								Regions	
p_{cor}	k_E	p_{unc}	p_{FWE}	p_{FDR}	T	Z	p_{unc}	X	Y	Z	L/R	Region
0	129	0	0.01	0.005	5.47	5.04	0	18	-78	-12	R	Occipital Cortex
			0.998	0.121	3.37	3.26	0.001	0	-81	-3		
0.037	54	0.004	0.398	0.041	4.31	4.09	0	-9	-21	48	L	Medial Frontal Gyrus
			0.609	0.057	4.1	3.91	0	-18	-12	51	L	Brodman Area 6
0.122	37	0.013	0.781	0.072	3.93	3.76	0	-30	-27	54	L	Precentral Gyrus
0.639	14	0.104	0.88	0.088	3.81	3.65	0	-33	15	21	L	Insula
0.76	11	0.146	0.891	0.089	3.8	3.64	0	-6	15	-3	L	Caudate head
0.72	12	0.13	0.947	0.095	3.69	3.54	0	15	-90	12	R	BrodmanArea18
0.72	12	0.13	0.982	0.104	3.57	3.43	0	33	6	21	R	Insula
			0.999	0.124	3.33	3.21	0.001	27	12	21		
0.562	16	0.084	0.997	0.12	3.41	3.29	0	-51	-66	6	L	Middle Temporal Gyrus
			0.999	0.124	3.33	3.21	0.001	-51	-57	3		

Table 8: *Gamble>Certain* regions: Local maxima of clusters, $p <$ 0.001 uncorrected, clusters with $k < 10$ voxels not shown (All local maxima uncorrected).

cluster			voxel								Regions	
p_{cor}	k_E	p_{unc}	p_{FWE}	p_{FDR}	T	Z	p_{unc}	X	Y	Z	L/R	Region
0.001	122	0	0.008	0.003	5.56	5.11	0	42	-24	60	R	Precentral Gyrus
0.012	72	0.001	0.008	0.003	5.56	5.11	0	-15	-75	9	L	Occipital Cortex

Table 9: *Certain>Gamble* regions: Local maxima of clusters, $p <$ 0.001 uncorrected, clusters with $k < 10$ voxels not shown (All local maxima uncorrected).

Cross-Correlation

Table 10 presents the cross-correlation of contrast values in our regions of interest. The OFC and amygdala contrast values are from Ambiguity>Risk contrast and striatum the opposite so a positive table entry between those areas is a negative correlation. These correlations show a modest link between amygdala and OFC, and a substantial (negative) correlation between striatum and OFC (again, the correlations shown are negative, but since contrasts are opposite, should be interpreted as positive). This is consistent with the

hypothesis that OFC is judging gradations of uncertainty and triggering differential striatal responses.

	R Amyg	L Amyg	R OFC	L OFC	R DMPFC	R DMPFC	R DStr	L DStr
R Amyg	-							
L Amyg	0.33 (0.21)	-						
R OFC	0.22 (0.42)	0.00 (0.99)	-					
L OFC	0.03 (0.91)	0.15 (0.58)	0.38 (0.14)	-				
R DMPFC	0.35 (0.18)	-0.10 (0.73)	0.76 (0.00)	0.58 (0.02)	-			
R DMPFC	0.23 (0.38)	0.12 (0.65)	0.46 (0.07)	0.09 (0.73)	0.56 (0.03)	-		
R DStr	-0.11 (0.69)	0.00 (1.00)	0.13 (0.63)	0.31 (0.24)	0.24 (0.37)	-0.31 (0.24)	-	
L DStr	-0.22 (0.40)	-0.16 (0.55)	0.14 (0.59)	0.61 (0.01)	0.08 (0.76)	-0.31 (0.24)	0.33 (0.22)	-

Table 10: Cross correlation (*p*-values) of pooled contrast values (each subject contributes one data point) between regions (ambiguity-risk contrast for amygdala and OFC; risk-ambiguity contrast for dorsal striatum). Note that since the contrasts are opposite for amygdala-OFC and striatum, positive correlations actually represents negative correlation in activity.

Correlation of fMRI and Behavioral Results

Stochastic choice analysis gives us a value of γ^j for each subject in the fMRI experiment. The γ_j estimated from the knowledge rounds positively correlates with brain activity across subjects as measured by the contrast values, $\beta^j_{amb} - \beta^j_{risk}$, between ambiguity and risk (averaged over the three treatments) in the right OFC (r = .55, p<.02, one-tailed), and left OFC (r =0.37, p< 0.1, one-tailed). This means that the subjects who exhibited more ambiguity-aversion revealed by choice, also show greater neural differences between risk and ambiguity in the OFC.

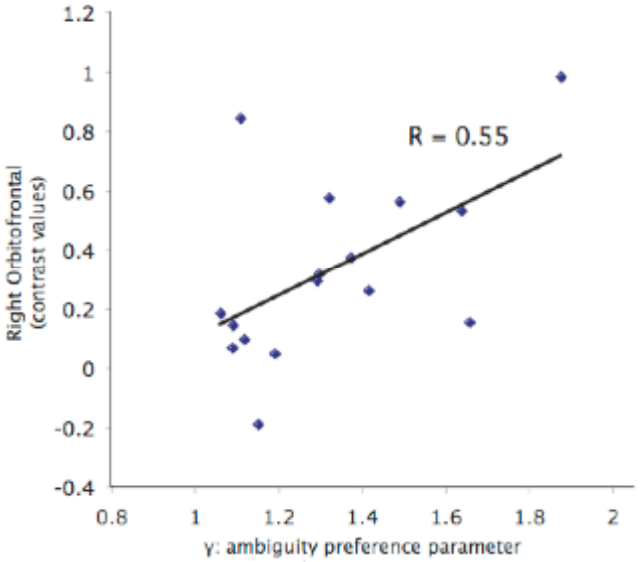


Figure 27: Correlation of behavioral measure of ambiguity aversion and contrast values. Contrast value is calculated as the difference between ambiguity and risk betas in the right orbitofrontal cortex.

Discussion

We present evidence that the human brain responds to varying levels of uncertainty, contrary to many decision theories which regard choices under risk and ambiguity as equivalent. FMRI data suggests that uncertainty is represented in a system that includes the amygdala and OFC.

The two hypothesized systems, amygdala/OFC, and striatum are active in both ambiguity and risk; the differences in activation are driven by the level of uncertainty in the different conditions. The fact that we see similar activation patterns for the real-world treatment as the card-deck treatment supports the hypothesis that risk and ambiguity are, in fact, points on a spectrum of uncertainty rather than two completely different entities. The reaction of the amygdala and OFC seems to be tied to the level of perceived uncertainty. The idea that ambiguity aversion in card deck and knowledge choices is related to rational aversion to betting against a better informed opponent (the informed opponent hypothesis) is supported

by similarities in time courses in the amygdala, OFC, and striatum between all three treatments.

An interesting implication of this study is that models of risk and ambiguity treating the two quantitatively, instead of qualitatively, different may be more neurally and, therefore, behaviorally accurate. The current models of risk aversion relying solely on the curvature of the utility function do not allow for this. The implication that both types of aversion are the result of a direct dampening of activity in the dorsal striatum, which may well be the internal representation of utility in the brain, could help resolve some of the paradoxes of risk aversion as well as ambiguity aversion: for example, the vastly different expressions of risk over small verses large bets.

A unified approach toward risk and ambiguity is also more appealing than a dual-systems model (separating risk and ambiguity) on evolutionary grounds and the available physiological evidence. Evolutionarily, it is difficult to imagine the adaptive advantages of possessing two (or more) systems for different types of uncertainty. This is especially so considering that in nature, as was mentioned at the outset of this chapter, known probabilities are exceedingly rare and most are limited to certain classes of games of chance. Physiologically, there is simply little evidence for such dual systems (see Sugrue et al. 2005 for a discussion of this topic).

Understanding the neural basis of choice under uncertainty, in the broader sense including both risk and ambiguity, is important because it is a fundamental activity at every societal level from retirement savings, to insurance pricing, to determining international military policy. These choices vary not only because of the presence of uncertainty, but the perceived level of uncertainty. Our results suggest that we pursue a unified model of uncertainty that would treat risk and ambiguity as points on a larger continuous scale. The knowledge treatment of the experiment further implies that the relevant level of uncertainty might be a function of mathematically unrelated factors, such as familiarity with related, but irrelevant, information.

Chapter 3

CAUSAL RELATIONSHIP BETWEEN BRAIN AND BEHAVIOR

Introduction

Although fMRI can teach us much about the neural correlates of perception and behavior, it is ultimately limited by its correlational nature. As reviewed on page 9, this can be overcome by utilizing other techniques that, unlike fMRI, are able to assess direction of causality. This chapter describes two studies assessing the direction of causality (both necessity and sufficiency) of the brain regions implicated in decision-making under risk and ambiguity in the previous chapter.

The first study uses behavioral data from lesion patients to examine the circuitry of choice under uncertainty. Lesions in the same areas activated in a scanning study can help us understand the interactions of the regions and establish whether a lesioned area is a necessary part of a hypothesized circuit. As described in Chapter 1, the lesion method is one of the oldest methods of studying the brain and still invaluable. The lesion method answers the question: is area X necessary for the expression of behavior Y. s

The second study uses fear conditioning with electrical stimulation in combination with choice under uncertainty. We use fear conditioning as an indirect method of stimulating the amygdala. Fear conditioning, using a wide variety of aversive stimuli (e.g., electrical stimulation, noise, taste), has been well established since the time of Pavlov and Skinner. More recently, it has been shown that certain brain areas, most importantly the amygdala, are critical to the acquisition and expression of fear conditioning (e.g., Selden et al. 1991; Labar et al. 1995; Quirk et al. 1995; Quirk et al. 1997).

The choice of fear conditioning is made because of its known robust and rapid activation of the amygdala (reviewed on page 64), and also because of technical limitations of current

techniques. The location of the amygdala, deep in the medial temporal lobe, precludes it from being affected by techniques such as TMS. Focal lesions of the amygdala are also extremely rare, and are commonly damaged along with other structures in the medial temporal lobe, such as the hippocampus. We therefore resort to using the indirect method of aversive conditioning with electrical stimulation to exogenously stimulate the amygdala.

Study 1: Orbitofrontal Lesion Patients

Background

The OFC is a stretch of cortex situated in the ventral (bottom) surface of the brain, lying on top of the orbit bone. This is one of the areas suffering damage in Phineas Gage, a railroad worker who famously had an iron rod blasted through his head but miraculously survived with little impairment to his intelligence or memory. What did seem to be affected, however, were his social skills and decision-making abilities (Harlow 1869; Ongur and Price 2000). This finding has been confirmed in other OFC patients over the past 80 years or so.

One task that proved useful in diagnosing such patients is the "Iowa Gambling Task (IGT)" (Bechara et al. 1996). The task consists of asking the subject having to choose between two decks of cards for say, 100 trials. Deck I is composed of high returns and high losses, Deck II low returns and low losses. Overall Deck II dominates I in both mean and variance. The general finding of the task is that whereas most control subjects learn to choose only from Deck II, OFC patients persist in choosing from Deck I. It is important to note here that subjects are not told the composition of the decks, and the sampling is done without replacement, so the deck's distribution is ambiguous. Clearly, the repeated nature (with feedback) of the experiment is a departure from the type of ambiguity present in the Ellsberg paradox. Subjects may hold different beliefs concerning the stationarity of the process.

In addition to behavioral changes, Bechara et al. also found that OFC patients do not exhibit normal GSR (Bechara et al. 1996). Normal subjects display heightened GSR during later trials when choosing the high-risk deck, upon which they cease to choose the high-risk

deck. OFC patients, on the other hand, keep choosing from the high-risk deck, and do not exhibit heightened GSR. This appears to be a global level effect, as these patients also have abnormally low responses to disturbing or exciting images.

Because of the extent of damage in a number of the patients, it is difficult to ascertain which substructure in the OFC is responsible for the behavioral changes. Neuroimaging and psychophysiological studies have given some clues to this. In a study using a gambling task, Patterson et al. also found that galvanic skin response (GSR) correlated with activity near OFC/Insula (Patterson et al. 2002).

Rogers et al. (1999) tested OFC patients, other prefrontal lesion patients, amphetamine abusers (a drug known to affect the dopaminergic system), normal subjects who were administered a drug that lowers plasma tryptophan levels (thereby simulating the effects of amphetamine), and normal controls. They used a variation of the gambling task with a known probability distribution. There was also a bet during each trial that the subject could take, allowing the experimenter to assess the risk attitudes of the subjects. They found that card choices of OFC patients, amphetamine patients, and normal subjects with lowered plasma tryptophan were similar. Even though the probability distribution is known, these subjects chose the color with the lower probability (a dominated choice) significantly more frequently than normal controls. OFC lesion patients, however, do not appear to be more risk averse; when allowed to wager on the outcome, lesion patients wagered significantly higher amounts than did normal controls.

Methods

Patient Background. Twelve neurological patients with single, focal, stable, chronic lesions of the brain were chosen from the Iowa Cognitive Neuroscience Patient Registry such that they were similar in terms of the etiology of their lesions (surgical resection), and similar on background neuropsychological measures.

We partitioned the patients on the basis of whether or not their lesion overlapped with the largest and most significant frontal activation focus we found in the fMRI study (right

lateral OFC, cf. Figure 19 and Table 4 in Chapter 2), or whether their lesion did not overlap with any significantly activated region.

These two groups, designated "frontal" and "control", consisted, respectively, of 5 (3 males, 2 females) and 7 patients (4 females, 3 males). The frontal group had lesions in bilateral OFC and frontal pole (1 patient), right OFC and right insula (1 patient), right OFC and frontal pole (1 patient) or only right OFC (2 patients) resulting from neurosurgical resection of brain tumors (frontal meningeoma resection). The control group had lesions in left (3 patients) or right (3 patients) anterolateral temporal cortex, or in left posterior temporal cortex (1 patient) resulting from neurosurgical resection for the treatment of epilepsy.

	OFC	*Comparison*	*t-statistic*
Age	54 (12)	52 (9)	0.31
VIQ	110 (21)	100 (9)	1.00
PIQ	117 (11)	100 (14)	2.35
FSIQ	114 (17)	100 (10)	1.65
MATH	102 (10)	98 (9)	0.71
WCST	6 (0)	4.9 (2)	1.45
MEMORY	106 (7)	100 (12)	1.09

Table 11: Lesion patient performance measures: Means (standard deviations) of VIQ, PIQ, FSIQ: verbal performance and full scale IQ from the Wechsler Adult Intelligence Test III or Revised. MATH: from the WRAT-R arithmetic subtest. WCST: Wisconsin card sorting test (number of categories successfully sorted). MEMORY: from the Wechsler Memory Scale III, general memory index.

There were no significant differences between these groups in the overall size of the lesion. All subjects had IQs in the normal range, had normal memory performance and arithmetic abilities, and were not aphasic, depressed, or perseverative (Table 11). Frontal-damage patients were not significantly different than temporal-damage controls except on the PIQ test, where frontal patients scored higher than temporal controls ($p < .05$).

Behavioral Task. In the ambiguity condition, patients were shown an actual card deck with 20 cards, in some mixture of red and black they could not see. They were given a series of choices between certain amounts of points (15, 60, 30, 40, 25 in that order) and bet on the color of their choice from the card deck for 100 points.

To illustrate, an ambiguity- and risk-neutral person would choose from the deck rather than take the certain amounts 15, 30, 40, or 25, but would take the certain 60 rather than choose from the deck. In the risky condition they were shown a deck with exactly 10 red and 10 black cards whose colors they could see. They made choices between a bet on the color of their choice from the deck for 100 points, or certain amounts of 30, 60, 15, 40, and 25 (in that order).

There are three small differences in this task and the Card-Deck treatment in the fMRI experiment reported in the previous chapter: (1) There were fewer choices in the lesion experiment, due to time constraints in conducting experiments with lesion patients and the need for multiple trials to extract fMRI signal; (2) there was wider range of certain point amounts in the lesion task (in case patients were extremely risk- and ambiguity-averse or preferring); and (3) due to human subjects restriction, the lesion task choices were not conducted for actual monetary payments.

Feature (1) means we could estimate ρ and γ for each individual in the fMRI study but were forced to pool data within each patient group for the OFC-control analysis. Note that the methodological differences (1-3) between the fMRI and lesion tasks do not matter for the most important finding from the lesion task, which is the significant difference between OFCs and controls.

Results

Aggregate Data. A summary of aggregate data is presented in Table 12. It is clear, through introspection, that the OFC patients choose the certain deck significantly less often than the control patients in both risky and ambiguous decisions. Moreover, whereas the control patients show a clear difference between ambiguous and risky decks, the OFC

patients appear to be indifferent between the two. This suggests that control patients are more ambiguity *and* risk averse than the OFC patients.

Lesion	*Certain Amt*	*Ambiguity*	*Risk*
Control	15	0.2857	0
	25	0.2857	0.1429
	30	0.5714	0.2857
	40	0.7143	0.5714
	60	0.7143	0.8571
OFC	15	0	0
	25	0	0
	30	0	0
	40	0.2	0.2
	60	0.4	0.6

Table 12: Summary of lesion patient choices. Proportion of patients choosing certain amount x when choosing between a gamble for 0 or 100 points versus x.

Stochastic Choice Model. To formalize this, we estimated a stochastic choice model similar to the one we estimated earlier in the fMRI study. Data from all subjects were pooled within each group. This yielded 25 choices (from 5 subjects) for the frontal group, and 35 choices (from 7 subjects) for the control group. The maximum likelihood estimation procedures are identical to those discussed on page 39.

To derive confidence intervals, a bootstrap procedure was used. We chose to use bootstrap instead of using the inverse of the Hessian matrix to derive standard errors. This is because the bootstrap estimate allows us to assess the bias of the ML estimate, which is an indication of shortcomings in the pooling assumptions that we made. In the fMRI experiment, all estimates were made *within* subjects, so that no pooling assumption is made.

The bootstrap consisted of 100 runs: For k-subject groups, a pseudo-sample consisted of k draws of different subjects with replacement from the sample. The estimation procedure was then applied to that pseudo-sample. One hundred pseudo-samples were drawn for each group.

The estimates for lesion patients using the data were (γ = 0.82, ρ = 1.09, λ = 0.10) for frontal patients, and (γ = 1.23, ρ = 0.74, λ = 0.27) for lesion controls. Across the 100 pseudo-samples, the bootstrap estimates yielded estimated means (standard deviations) of γ = 0.88 (0.28), ρ = 1.15(0.20), λ = 0.48(1.29) for the frontal patients, and γ = 1.25 (0.22), ρ = 0.77 (0.11), λ = 0.28 (0.16) for the lesion patients.

The only noticeable difference in the pooled estimates and the central tendencies of the bootstrapped estimates, is the lower λ for frontals (.10) in the group procedure, versus, the bootstrap mean (0.48). This reflects the fact that one subject behaved differently than the others and more randomly across certain-amount x values. Fitting group data including this subject requires estimation of a relatively low λ. However, in bootstrap runs this outlying subject was often not selected into the pseudosample, and in those pseudosamples more inflected responses to x deliver higher λ. The crucial parameter, the ambiguous probability discount γ, is quite close in the two types of estimation however, and is close to the ambiguity-neutral value of 1.

The two-dimensional 90% confidence interval was calculated from the bootstrap by calculating the minimum ellipsoid that contains 90% of the pseudosample estimates (Fox 1997). The angle of the ellipse reflects correlation between the parameters. Figure 23 shows 90% confidence ellipsoids for the frontal and control patients for the (γ, ρ) parameter pairs. Frontal patients are risk-and ambiguity-neutral, i.e., the hypothesis that $\gamma = \rho = 1$ cannot be rejected. This behavior of frontal patients was significantly different than the damage control group, who were averse to both risk and ambiguity. The OFC-lesioned group was therefore abnormally neutral toward ambiguity (which is, ironically, a hallmark of rationality under SEU).

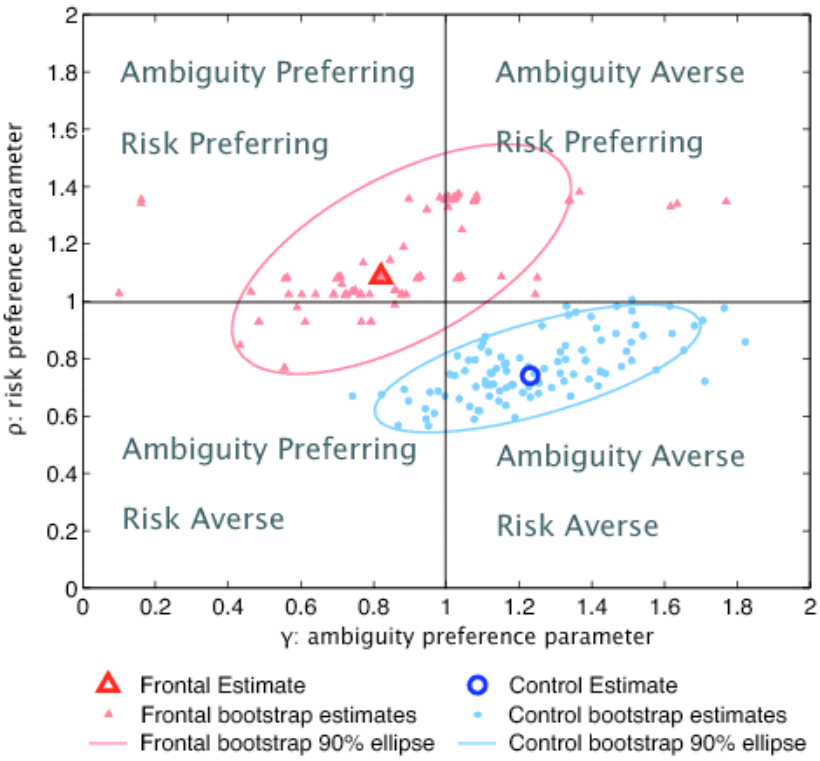


Figure 23: Measures of risk (ρ) and ambiguity (γ) preferences of OFC (n =5) and control group (n =7). The risk neutral line (γ = 1) and the ambiguity neutral line (p = 1) demarcate four quadrants as labeled. Open symbols plot ML estimates of a group-level stochastic choice model (frontals: (γ =0.82, ρ =1.09); lesion controls: (γ =1.23, ρ = 0.74)). Solid symbols represents 100 bootstrapped (γ,ρ) estimates. Correlation between ρ and γ (tilt of ellipse) is 0.42 for frontal and 0.31 for control.

Linking Behavior, Neuroimaging, and Lesion Data. The addition of the lesion data allows us to do one more interesting calibration exercise. In particular, the parameter γ enables us to link the fMRI and lesion studies. Assume that the frontal patients would have a right OFC (ROFC) contrast value of zero if they were imaged during these tasks (since all have ROFC damage). Then we can guess what value of γ the OFC patients might exhibit behaviorally, by extrapolating correlation between ROFC activity and γ in the case where there is zero activity in ROFC. This extrapolation gives a predicted γ = 0.85. The actual value estimated from the OFC patients' behavioral choices is γ = 0.82, which is reasonably close to the extrapolated prediction.

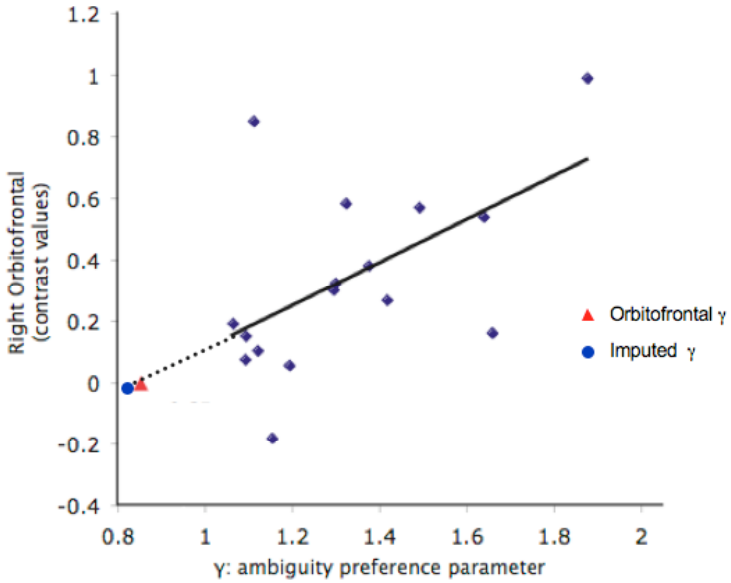


Figure 24: Imputed ambiguity preference parameter (γ) for OFC lesion patients. Imputed value was calculated from regression model using neuroimaging and behavioral data from normal subjects (see Figure 23).

Discussion

The regions implicated in our fMRI experiments and confirmed by behavioral experiments with lesion patients have been observed in previous studies using different tasks. The striatum-amygdala-OFC network is well-established in animal and human studies as a system for reward learning, including probabilistic learning (Critchley et al. 2001). The OFC is highly interconnected with the basolateral amygdala. These interconnections appear to play vital roles in learning and reversal learning in rats (Schoenbaum et al. 2000).

Lateral OFC, in particular, appears to be necessary to change existing associations (O'Doherty et al. 2003). Our finding that the OFC is activated as a function of ambiguity, and that its damage reduces sensitivity to ambiguity, suggest that this structure is a necessary component for reacting to gradations of uncertainty.

This hypothesis fits well with current knowledge of amygdala and OFC function. Both the amygdala and OFC are known to receive rapid, multi-modal sensory input; both are

bidirectionally connected and known to function together in evaluating the value of stimuli (Gaffan et al. 1993), and both are likely involved in the detection of salient, relevant, and ambiguous stimuli. The latter function has been hypothesized especially for the amygdala (Whalen 1998; Adams et al. 2003). Critically, such a function also provides a reward-related signal that can motivate behavior, in virtue of the known connections between the amygdala/OFC and the striatum (Amaral et al. 1992). Under ambiguity, the brain is alerted to the fact that information is missing, that choices based on the information available, therefore, carry more unknown (and potentially dangerous) consequences, and that cognitive and behavioral resources must be mobilized in order to seek out additional information from the environment.

Study 2: Choice Behavior under Fear Conditioning

Background

The amygdala is a complex of nuclei situated at the end of the hippocampus in the medial-temporal lobe (MTL). Amygdala neurons respond to primary rewards and reward predicting stimuli (Schultz et al. 2000). Its role in emotions and conditioning is well established, especially in the domain of fear. It is well documented that amygdala activity is not mediated by consciousness. In particular, the subliminal priming literature has shown convincingly that the amygdala will respond to emotional faces even though they are not consciously seen. The amygdala is crucially involved in learning (O'Doherty 2003). Lesioning of the amygdala in rats impairs associative learning. It is hypothesized on the basis of the strong reciprocal connections between the amygdala and the OFC, and a massive input from dopamine neurons, that the amygdala-OFC circuit is crucial in learning.

In the past two decades, it has been established that the amygdala is critical in the acquisition and expression of fear conditioning. Labar et al. (1995) showed that human subjects with damage to the amygdala showed typical autonomic responses (as measured by galvanic skin response) to mild electrical shocks but did not respond to the conditioned stimulus. Quirk et al. (1995) recorded from the lateral nucleus of the amygdala in free

behaving rats, and showed that neurons in the lateral nucleus responded to tones that signaled the onset of electrical shocks during acquisition.

The fear-conditioning response serves a vital function in survival of organisms. The amygdala has long been known to be involved in the acquisition of fear conditioning, including evidence, among other species, rats (Garcia et al. 1999) and humans (Whalen et al. 1998; Phelps et al. 2001)

Method

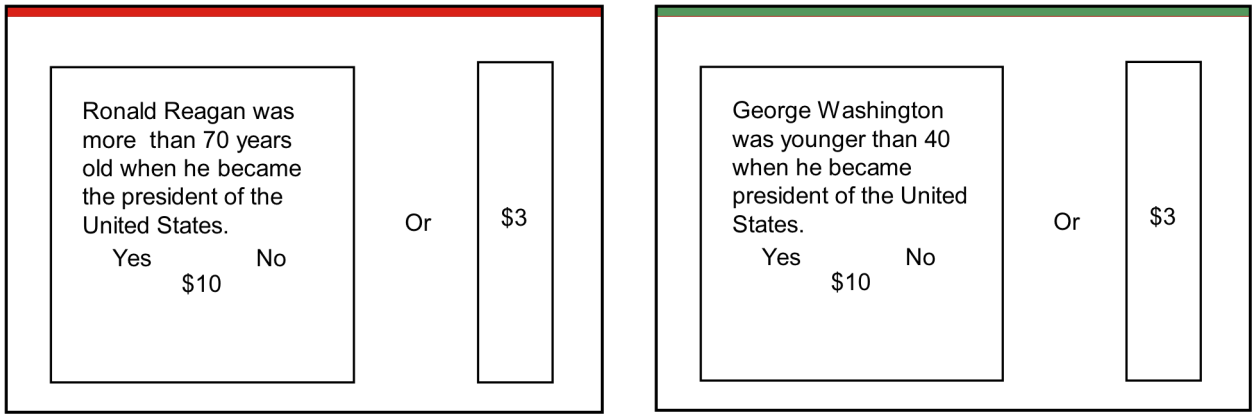
Subjects. Twenty-two Caltech undergraduate and summer research students (16 male, 6 female) completed the experiment. Informed consent was reviewed and approved by the Caltech committee for the protection of human subjects.

Electrical Stimulation. Conditioning stimuli were presented using equipment and software from Contact Precision Instruments (www.psylab.com). Silver/Silver Chloride electrodes filled with Med Associates paste TD-246 were used for shock presentation.

Shocking electrodes were attached to the palmar surface of the index and middle fingers of the left hand. The electrical stimulation used in these experiments was a 100ms long, constant 60 Hz AC shock. Each individual's shock level was determined using a subjective rating protocol that sought a level that was "uncomfortable, but not painful". This shock level was used throughout the experiment.

Behavioral Task. Subjects were presented with a series of 100 binary gambles in the form of true/false trivia questions. The form of the presentation was similar to the Knowledge Treatment described in Chapter 2. We used trivia instead of card decks because the effects from ambiguity aversion are larger in the Knowledge treatment than the Card-Deck treatment. The subject can choose to bet on the trivia question (with associated payoff x or 0 indicated below the question), or choose a certain outcome (with associated payoff indicated on the right hand side of the screen).

On top of the screen on each trial was a colored bar indicating whether the subject will be shocked following the choice—red indicates shock, green no shock. The shock is delivered following the decision of the subject, regardless of the alternative chosen by the subject. Finally, the trials that were chosen to be the shock trials were switched between subjects to counterbalance and ensure that there were no question effects (12 completed one version, and 10 counterbalanced version). The two forms of the Figure 25 shows an example of the two types of trials.



(A) Shock (B) No Shock

Figure 25: Example screens of the experiment. (A) Red bar on top of screen indicates shock; (B) Green bar on top of screen indicates no shock.

Results

Shock Calibration. Because levels of shock were determined by the subjective perceptions of the subjects, substantial variations exist in the calibration. Figure 26 shows a sorted barplot of the level of shocks calibrated for each subject.

There are two primary sources of variation: one perceptual and one cognitive. The perceptual difference arises from variance in pain tolerance. The cognitive aspect depends on how subjects interpret "uncomfortable, but not painful." We believe that most of the variation is due to the cognitive aspect.

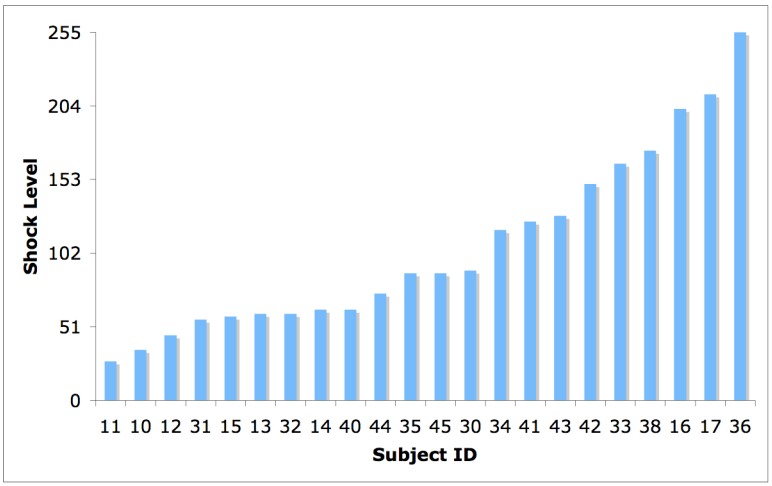


Figure 26: Sorted barplot of levels of administered shocks. Maximum level allowable by IRB is 255.

Accuracy. Of the trials where the gamble was chosen, mean (s.e.) accuracy were 0.510 (0.0307) in the no shock trials and 0.505 (0.0339) in the shock trials. Subject accuracy did not differ significantly between the shock and no shock conditions. A t-test of the difference between the two treatments returned a non-significant *p*-value of 0.796, and a *p*-value of 0.86 in a weighted regression using the number of choices as the weights (there was substantial variability in subjects' risk aversion coefficients, and thus the number of times that they chose the certain outcome).

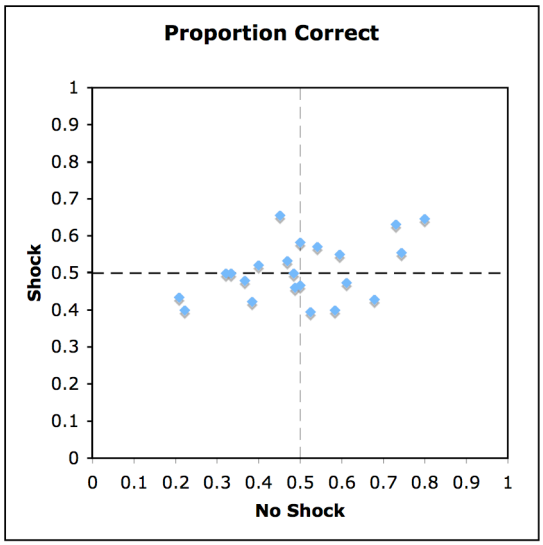


Figure 27: Scatterplot of proportion of correct guesses per subject in shock condition versus no shock condition.

Figure 27 shows a scatterplot of the proportion of correct guesses in the shock and no shock conditions. It is clear from the plot that although the mean accuracy is similar between the treatments, the variance is much larger in the un-shocked condition. This is significant in a Bartlett test for homogeneity of variance at $p < 0.002353$ (Bartlett's K-squared = 9.2513, df = 1). In addition, the difference in variance did not vary significantly with the level of shock. The reason behind this difference, however, is unclear.

Choice Data. Table 13 shows the summary of the subjects' behavioral data. Included are the levels of shock that were calibrated for each subject, as well as the number of times the subject made each choice. Differences between the proportion of choosing the certain choice in the shock and no shock conditions suggest differences in risk/ambiguity attitudes of subjects.

		Proportion Certain	
Subject	**Shock Level**	**No Shock**	**Shock**
10	35	0.44	0.52
11	27	0.44	0.3
12	45	0.52	0.54
13	60	0.36	0.4
14	63	0.28	0.3
15	58	0.22	0.28
16	202	0.48	0.62
17	212	0.34	0.48
30	90	0.52	0.52
31	56	0.64	0.62
32	60	0.74	0.48
33	164	0.82	0.8
34	118	0.52	0.44
35	88	0.22	0.22
36	255	0.16	0.2
38	173	0.52	0.52
40	63	0.38	0.36
41	124	0.76	0.7
42	150	0.7	0.66
43	128	0.4	0.5
44	74	0.2	0.14
45	88	0.5	0.54

Table 13: Summary of subject results, including shock level calibrated and the proportion of certain choices that the subject chose (out of 50) in No Shock and Shock conditions.

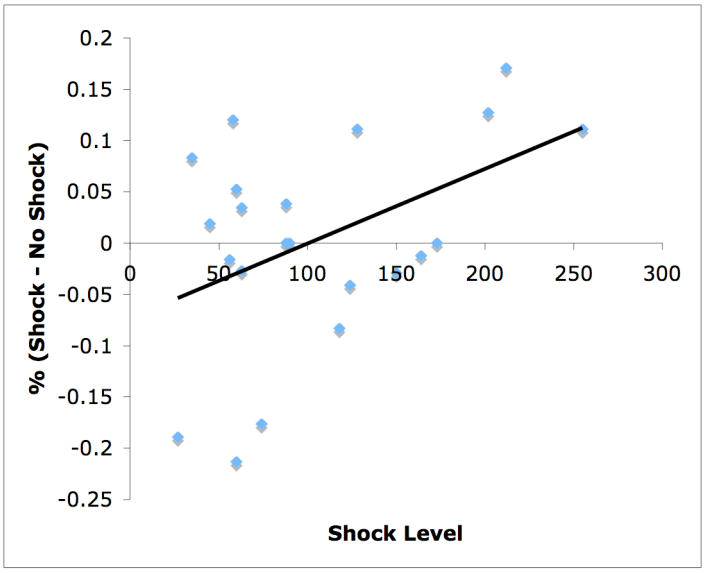


Figure 28: Scatterplot of percent difference between Shock and No Shock conditions versus the shock level that subjects received.

Figure 28 shows the percent difference in choosing the certain choice ((*# Certain | Shock – # Certain | No Shock) / # Certain*) versus the level of shock that was delivered to subjects. Regression analysis revealed that both the intercept and shock interaction coefficients were significant.

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	-0.07358845	0.039872403	-1.845598582	0.079808221
Shock Level	0.000728285	0.000325329	2.238612483	0.036712743

$R^2 = 0.200$, *F-Statistic* = 5.011, *Significant F* = 0.0367

Tests for nonlinearity in the form of quadratic relationship was negative.

```
                Estimate   Std. Error   t value  Pr(>|t|)
(Intercept)     -0.00149   0.0263       -0.054   0.958
shock           0.172      0.0984       1.746    0.109
shock^2         0.010      0.0984       1.014    0.332

Residual standard error: 0.09839 on 19 degrees of freedom
Multiple R-Squared: 0.2704, Adjusted R-squared: 0.1378
F-statistic: 2.039 on 2 and 11 DF,  p-value: 0.1765
```

Similarly, robust regression showed nearly identical results as OLS.

```
               Value  Std. Error  t value
(Intercept)  -0.0559   0.0449    -1.2460
shock         0.0007   0.0004     1.6947
```

This is likely due to the fact that there is no significant heteroskedasticity in the data. A Breusch-Pagan test against heteroskedasticity is not significant (BP = 0.719, df = 1, p-value = 0.396). Figure 29 shows the Q-Q plot of the residuals for the OLS regression.

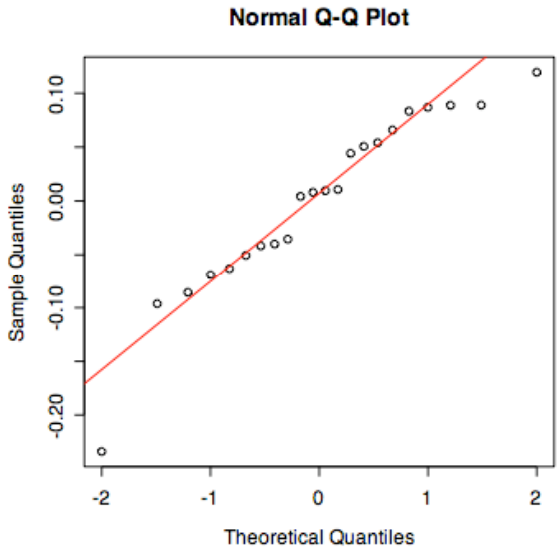


Figure 29: Q-Q plot of empirical and theoretical residuals.

The significant effect of the shock level is consistent with our initial hypothesis that fear conditioning increases risk/ambiguity aversion. The significant (albeit weaker) intercept, however, was not part of our initial hypothesis. This suggests that weak electrical stimulation, in fact, lowers the degree of risk/ambiguity aversion. This may be explained partially by the fact that some of the shock levels are extremely low, and that it may be moving subjects along the Yerkes-Dodson curve up until it crosses some subjective threshold level.

Stochastic Choice Model. To formalize this, two stochastic choice models were estimated. The first is identical to those described on page 39. This model focused on the parameter γ that operates on the subjective probability function. The difference with the second model

is that differences between the conditions are captured in the utility function rather than the probabilities. That is, $u_{noshock} = x^{\rho}$ in the no shock condition and $u_{shock} = x^{\rho+\rho_{shock}}$ in the shock condition. This is to capture the possibility that the fear conditioning is operating on the curvature of the utility function rather than the decision weights.

	Model 1			*Model 2*			
Subject	γ	ρ	λ	ρ_{shock}	ρ	λ	*Shock*
10	1.261 (0.323)	1.089 (0.188)	0.119 (0.125)	-0.139 (0.208)	1.034 (0.131)	0.185 (0.131)	35
11	0.873 (0.073)	0.955 (0.055)	0.613 (0.19)	0.109 (0.068)	0.964 (0.061)	0.483 (0.15)	27
12	*	*	*	0.109 (0.068)	0.964 (0.061)	0.483 (0.15)	45
13	1.025 (0.062)	1.009 (0.045)	0.682 (0.193)	-0.029 (0.052)	1.011 (0.04)	0.714 (0.183)	60
14	1.033 (0.062)	1.096 (0.05)	0.519 (0.16)	-0.008 (0.048)	1.083 (0.044)	0.549 (0.153)	63
15	1.027 (0.069)	1.145 (0.059)	0.377 (0.13)	-0.036 (0.052)	1.147 (0.05)	0.4 (0.126)	58
16	1.118 (0.092)	0.917 (0.052)	0.725 (0.221)	-0.127 (0.084)	0.923 (0.043)	0.89 (0.217)	202
17	1.336 (0.206)	1.234 (0.141)	0.13 (0.099)	-0.127 (0.114)	1.136 (0.102)	0.21 (0.121)	212
30	0.984 (0.047)	0.881 (0.031)	1.674 (0.358)	-0.032 (0.043)	0.903 (0.028)	1.644 (0.33)	90
31	0.84 (0.709)	0.129 (0.405)	0.928 (0.803)	-0.135 (0.419)	0.224 (0.496)	0.822 (0.659)	56
32	0.635 (0.178)	0.582 (0.152)	1.034 (0.354)	0.371 (0.188)	0.547 (0.184)	0.639 (0.196)	60
33	0.883 (0.42)	0.012 (0.224)	2.97 (1.336)	-0.091 (0.189)	0.064 (0.255)	2.811 (1.206)	164
34	0.917 (0.113)	0.935 (0.085)	0.407 (0.175)	0.106 (0.112)	0.92 (0.098)	0.352 (0.142)	118
35	1.008 (0.09)	1.231 (0.087)	0.208 (0.097)	0.014 (0.06)	1.218 (0.076)	0.213 (0.091)	88
36	1.054 (0.054)	1.164 (0.047)	0.576 (0.167)	0.008 (0.042)	1.129 (0.041)	0.634 (0.168)	255
38	0.973 (0.088)	0.877 (0.061)	0.698 (0.219)	0.023 (0.086)	0.878 (0.064)	0.666 (0.187)	173
40	1.005 (0.067)	1.022 (0.05)	0.588 (0.179)	0.04 (0.055)	0.997 (0.048)	0.603 (0.166)	63
41	0.855 (0.207)	0.447 (0.173)	1.273 (0.465)	-0.011 (0.301)	0.478 (0.185)	1.081 (0.33)	124
42	0.756 (0.4)	0.182 (0.308)	1.285 (0.774)	-0.342 (0.328)	0.388 (0.321)	1.067 (0.519)	150
43	1.123 (0.209)	1.093 (0.154)	0.138 (0.118)	-0.182 (0.214)	1.091 (0.108)	0.185 (0.121)	128
44	0.921 (0.051)	1.107 (0.042)	0.657 (0.183)	0.021 (0.044)	1.142 (0.041)	0.544 (0.15)	74
45	1.248 (1.042)	0.812 (0.443)	0.097 (0.186)	-0.072 (0.681)	0.791 (0.519)	0.137 (0.151)	88

Table 14: Stochastic choice estimates of subjects and associated

shock levels for the two models. (*: parameters for subject 12 were poorly identified in model 1).

Table 14 summarizes the stochastic choice models of the subjects. Figure 30 shows scatterplot of the stochastic choice model estimate of the parameter ρ_{shock} and shock level. The first model did not yield significant results and is not displayed. This stochastic choice model was estimated by assuming that subjective probability level for all questions is 0.5. Another model was estimated using the mean of the subject's accuracy. The results are similar to the first model. Finally, a third model was estimated using the mean accuracy of each question. In this case, however, some questions were rarely answered and therefore had extreme probabilities nearly 0 and 1, leading to identification issues.

A weighted regression using the precision parameter λ (which measures the noisiness of the subject's utility function) as the weight revealed a significant effect on the shock parameter at $p < 0.1$.

```
Estimate    Std. Error t value     Pr(>|t|)
(Intercept)  0.0631845  0.0624678   1.011   0.3239
shock       -0.0008718  0.0004748  -1.836   0.0813
--
Multiple R-Squared: 0.1442,     Adjusted R-squared: 0.1014
F-statistic: 3.371 on 1 and 20 DF,  p-value: 0.08128
```

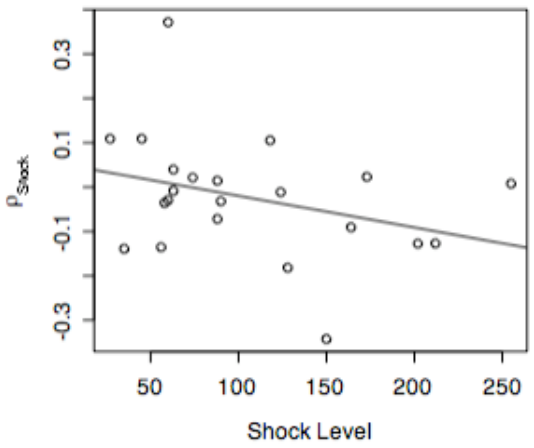


Figure 30: Scatterplot of stochastic choice model parameter ρ_{shock} and shock level.

Residuals are not fat-tailed but are correlated with the precision parameter (lambda). There is no evidence of heteroskedasticity in the data. A studentized Breusch-Pagan test failed to reject the homoskedastic null hypothesis (BF = 0.0725, df = 1, $p < 0.788$). Figure 31 shows a Q-Q Plot of Base model (OLS with no quadratic).

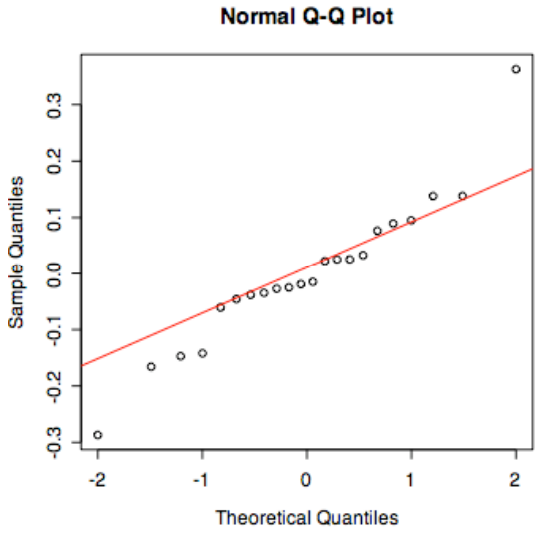


Figure 31: Q-Q Plot of the base model (OLS with no quadratic term).

Moreover, adding a quadratic term does not significantly improve the fit versus the OLS or the WLS.

```
               Estimate  Std. Error   t value   Pr(>|t|)
(Intercept)    -0.0241    0.0284      -0.847     0.408
shock          -0.2058    0.1335      -1.541     0.140
shock^2         0.1365    0.1335       1.023     0.319
---
Residual standard error: 0.1335 on 19 degrees of freedom
Multiple R-Squared: 0.152, Adjusted R-squared: 0.0634
F-statistic: 1.71 on 2 and 19 DF,  p-value: 0.207
```

WLS with the quadratic fits the data nearly as well as that of WLS without the quadratic.

```
               Estimate  Std. Error  t value   Pr(>|t|)
(Intercept)    -0.0204    0.0281     -0.727     0.4761
shock          -0.2416    0.1331     -1.815     0.0854
shock^2         0.1972    0.1335      1.478     0.1558
---
Residual standard error: 0.1055 on 19 degrees of freedom
```

```
Multiple R-Squared: 0.232, Adjusted R-squared: 0.152
F-statistic: 2.877 on 2 and 19 DF,  p-value: 0.0810
```

An F-test confirms that we cannot reject the WLS model without the quadratic term for the full model.

```
Analysis of Variance Table
Model 1: ρshock ~ shock
Model 2: ρshock ~ shock + shock^2
  Res.Df      RSS     Df    Sum of Sq    F        Pr(>F)
1     20      0.235
2     19      0.211   1     0.024        2.1845   0.1558
```

Discussion

This study yielded mixed results about the effects of fear conditioning on choice under uncertainty. Mild electrical stimulation appears to have a positive relationship with the degree of risk aversion. Low levels of stimulation, however, appear to actually make subjects *more* risk loving. The most puzzling finding, however, is that electrical stimulation appears to reduce the variance in the accuracy of the subjects' responses, without affecting the mean accuracy.

It is unclear why a number of subjects were guessing at less than chance. This does not appear to be due to questions since questions were counterbalanced. Table 15 shows that there is no correlation between accuracy and shock level.

Treatment	*Pearson's r*	*p-value*
Accuracy|Shock	0.37	0.19
Accuracy|No Shock	0.05	0.86
Accuracy|Shock / Accuracy|No Shock	0.02	0.94
(Accuracy|Shock – Accuracy|Shock)/ Accuracy|Shock	0.15	0.60

Table 15: Correlation of subject's various measures of accuracy and shock level. (Accuracy|Shock: Accuracy in Shock condition, Accuracy|No Shock: Accuracy in the No Shock condition)

This study suffers from limitations that would serve to strengthen the robustness of the result. As mentioned before, the most difficult problem encountered was the substantial variation in levels of shock that are administered to subjects. There is good reason to believe that much of this variation is not due to differences in the perception of pain, but

rather differences in interpretation of the statement "uncomfortable, but not painful". Obtaining measures such as galvanic skin response can, to some degree, alleviate this problem, but there, too, exists much individual difference.

The fact that the regression is significant in the second model but not the first suggests that the shocks are operating directly on the curvature of the utility function rather than the decision weights. This is entirely possible due to the known connections between the amygdala and striatum (as well as the OFC and the striatum). A possible way to test this hypothesis would be to image subjects making decisions while being paired with shocks. Other possibilities include giving objective probabilities as opposed to having to infer or impose subjective probabilities in the lotteries, thereby separating the effects on probabilities and utility function.

Conclusion

This chapter attempts to understand the neural causes of decision-making under uncertainty. Two techniques—the lesion method and fear conditioning—were used to assess necessity and sufficiency, respectively in separates studies. Taken as a whole, these results appear to validate the hypothesis that the brain regions implicated in the previous chapter (OFC and amygdala) are part of a circuit central to decision-making under uncertainty.

This leaves the striatum as the only key area of our circuit unexplored. The striatum, however, is perhaps the least controversial region in this circuit. All our prior knowledge in reward anticipation and reward learning suggests the striatum reflects the reward value. It is not surprising, therefore, that the striatum is less activated under ambiguity than risk, since behaviorally, decisions under risk are preferred to the formally identical decisions under ambiguity. Possible studies of the role of the striatum include using patients who have degenerative diseases affecting the striatum.

Recent work on decision-making in Parkinson's disease, which is marked by a degeneration of the dopaminergic system, including the caudate, shows that Parkinson's

patients are impaired in the Iowa Gambling Task (Mimura et al. 2006). Furthermore, work by Pessiglione et al. (2005) suggests that stimulation of the basal ganglia (of which the striatum is a subset) is not sufficient to promote proper functioning in Parkinson's patients; dopaminergic medication, rather than deep brain stimulation, was able to correct behavioral interference in a decision task.

Taken as a whole, these results confirm the role of the dopaminergic system in expected reward valuation and learning. In addition, it suggests that this system takes into account information that would be disregarded by standard decision theory.

Chapter 4

PROBABILITY WEIGHTING FUNCTION IN THE BRAIN

Introduction

One of the central tenets in classical decision theory is that expected utility is linear in probabilities (Savage 1954). Therefore, although the utility function is allowed to be nonlinear (and in particular, concave), the decision-weight is not.

This is a consequence of the independence axiom (and its variants, such as the sure-thing principle) that underlies expected utility theory (EUT). The independence axiom states that a decision-maker's preference ordering should not be changed with the addition of a common consequence. That is, $x \succ y \Rightarrow ax + (1-a)z \succ ay + (1-a)z, \forall a \in (0,1)$.

That people might not weight probabilities linearly, as dictated by SEU was first brought to attention by Maurice Allais. Allais proposed the Common-Ratio Effect, or Allais's Paradox (Allais 1953). For example, a decision-maker may prefer a sure gain of $100,000 over a coin toss for $300,000 or nothing, but at the same time prefer a 1% chance for $300,000 over a 2% chance for $100,000.

As noted previously by Prelec (1998), this example brings many of the same intuitions invoked when discussing the concept of diminishing marginal utility. However, whereas it is uncontroversial to state that the utility difference between $0 and $50 is greater than that between $100,000 and $100,050, no such nonlinearities are admitted in the decision weights that people use to arrive at expected utilities.

The existence of nonlinear weighting of probabilities has obvious impact for the financial (Barberis and Huang 2004) and gambling (Alistair C Bruce 2000) markets, but it also has profound implications for health care (Bleichrodt et al. 1999) and criminal justice

(Lattimore et al. 1992). The nonlinear hypothesis also needs to be separated from a skewness-preference hypothesis, which has been invoked previously to account for apparent risk seeking behavior (Kraus 1976; Golec 1998; Garrett and Sobel 1999). A skewness-preference explanation does not imply a violation of SEU, whereas a nonlinear weighting explanation does. To the extent that these models predict different behavioral implications, they also suggest different policy recommendations.

Literature Review

Despite the well-documented nature of this phenomenon, a psychological interpretation was lacking until it was introduced by Tversky and Kahneman in 1992. This interpretation would be familiar to any psychophysicist. Tversky and Kahneman (1992) proposes that the properties of the weighting function for their Cumulative Prospect Theory (CPT):

> *"…reflect the principle of diminishing sensitivity: the impact of a change diminishes with the distance from the reference point… In the evaluation of uncertainty, there are two natural boundaries—certainty and impossibility—that correspond to the endpoints of the certainty scale. Diminishing sensitivity entails that the impact of a given change in probability diminishes with its distances from the boundary… Diminishing sensitivity, therefore, gives rise to a weighting function that is concave near 0 and convex near 1." (p. 303)*

Note, however, that in CPT and rank-dependent utility (RDU) models in general, the decision weight is not equivalent to the probability weighting function. That is because the decision weight is further transformed depending on the rank of the outcomes. In CPT, the value of the highest outcome is attached weight of $w(p)$, and all other outcomes by

$$w\left(p_i + \sum_{j=1}^{i-1} p_j\right) - w\left(\sum_{j=1}^{i-1} p_j\right).$$

This is because, under CPT, given gambles $(p_1, x_1; \ldots; p_n, x_n)$, where $|x_i| > |x_{i+1}|$ and the x's are on the same side of the reference point, the expected utility of the gamble is

$$w(p_1)v(x_1) + \sum_{i=2}^{n}\left[w\left(\sum_{j=1}^{i} p_j\right) - w\left(\sum_{j=1}^{i-1} p_j\right)\right]v(x_i).$$

One interpretation is that the weighting function captures the psychophysics of the perception of probabilities, whereas the decision weight captures how decision-makers combine probabilities (or rather, the perceived probabilities) and outcomes (Wu and Gonzalez 1996). As the descriptive validity of CPT and RDU in general is an open question, we will sidestep this issue in this paper by considering only binary gambles with one outcome fixed to 0 and another strictly greater.

A psychological interpretation is important because Allais's paradox is, as noted above, a violation of Savage's SEU. Therefore, a decision maker who exhibits Allais's Paradox is also subject to arbitrage opportunities. A psychological interpretation can be seen as the link between the neural mechanism of the decision-making process and the behavior of the decision-maker.

Functional Forms

At least three functional forms have been suggested in this literature.

Lattimore Function

Gonzalez and Wu (1999) and Lattimore et al., (1992) assumed that the log odds ratio of the perceived probabilities are an affine transformation of the actual probabilities, such that

$$\log\frac{\pi}{1-\pi} = \gamma \log\frac{p}{1-p} + \tau.$$

Solving for π yields

$$\pi = \frac{\delta p^{\gamma}}{\delta p^{\gamma} + (1-p)^{\gamma}},$$

where $\delta = e^{\tau}$.

Gonzalez and Wu (1999) motivated this transformation by proposing two different properties of the weighting function. The curvature of the function is controlled by γ, which is the *discriminability* of stimulus. The second parameter δ accounts for the *attractiveness* of the gamble, which corresponds to the relative over/underweighting of probabilities in general by the decision-maker. Figure 32 shows the effects of the parameters on the shape of the probability weighting function.

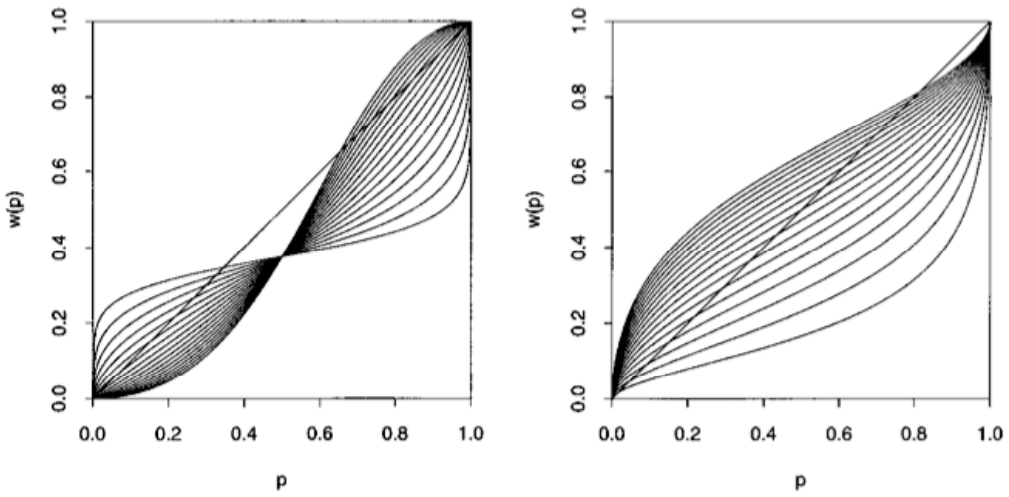


Figure 32: Effects of γ on the curvature and δ on the elevation of the probability weighting function. Left panel fixes δ = 0.6 and varies 0.2 < γ < 1.8. Right panel controls γ = 0.6 and varies 0.2 < δ < 1.8. Adapted from Gonzalez and Wu (1999).

Tversky and Kahneman

A similar one-parameter functional form was proposed by Tversky and Kahneman (1992). However, instead of allowing different weights of the probabilities, it takes an exponent of

the denominator. The function is $w(p) = \frac{p^{\gamma}}{(p^{\gamma} + (1-p)^{\gamma})^{1/\gamma}}$. This function exhibits an inverse S-shape for parameters between 0.27 and 1. The intuition for this function is similar to that of the Lattimore function.

Prelec

Prelec (1998) axiomatized a widely-used specification due to its axiomatic foundation and its parsimony. In a restricted version, it has only one parameter and is therefore widely used in empirical studies.

The Prelec function itself is

$$w(p) = \exp\left\{-\beta(-\ln p)^{\alpha}\right\}$$

where the one-parameter specification is

$$w(p) = \exp\left\{-(-\ln p)^{\alpha}\right\}.$$

The one-parameter version has the additional property that it intersects the diagonal at $1/e$ or ~0.36—a value that is near those found in past empirical studies.

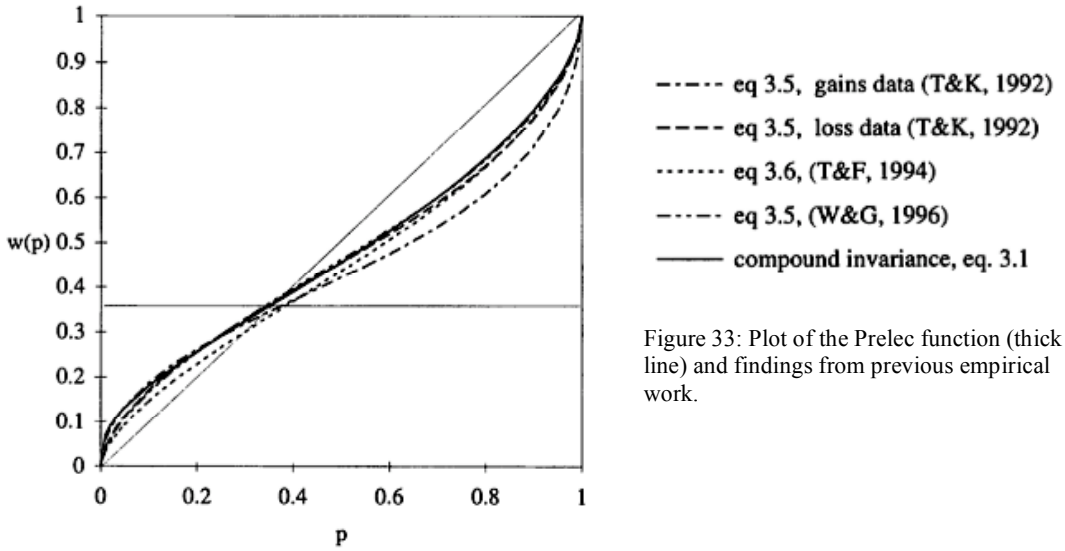


Figure 33: Plot of the Prelec function (thick line) and findings from previous empirical work.

The key property of the Prelec function is that of compound invariance.

Compound Invariance: Let N be any natural number. Then *N-compound invariance* holds iff, for consequences *x, y, x', y'*, probabilities *p, q, r, s* in (0,1), with $q < p$, $r < s$,

$$(x,p) \sim (y,q) \text{ and } (x,r) \sim (y,s) \text{ implies}$$

$$(x', p^N) \sim (y', q^N) \Rightarrow (x', r^N) \sim (y', s^N).$$

For example, a decision-maker may exhibit the Allais's paradox with the following pair of choices:

($10,000 for sure) ~ (1/2 chance of $30,000)

(1/2 chance of $10,000) ~ (1/6 chance of $30,000).

Then we can find an amount x such that

($*x* for sure) ~ (1/4 chance of $30,000)

(1/4 chance for $*x*) ~ (1/36 chance of $30,000).

The intuition for this axiom is that people should be consistent to changes in probability, or "compounding".

Empirical Literature on Probability Weighting

It is a testament to the robustness of this phenomenon that a wide variety of studies are in agreement concerning the general properties of the weighting function, using a variety of estimation techniques, elicitation methods, types of outcome, and functional forms.

Table 16 presents these studies and their results. I present only the studies using parametric forms, as it is difficult to compare and visualize nonparametric results (although they are in general agreement. in terms of qualitative results (e.g., Wu and Gonzalez 1996). Table 17 presents a more detailed table of these studies, including the elicitation and estimation

procedures, the stimuli used, as well as the findings[10]. Figure 34 plots some of the functional forms studied in the literature.

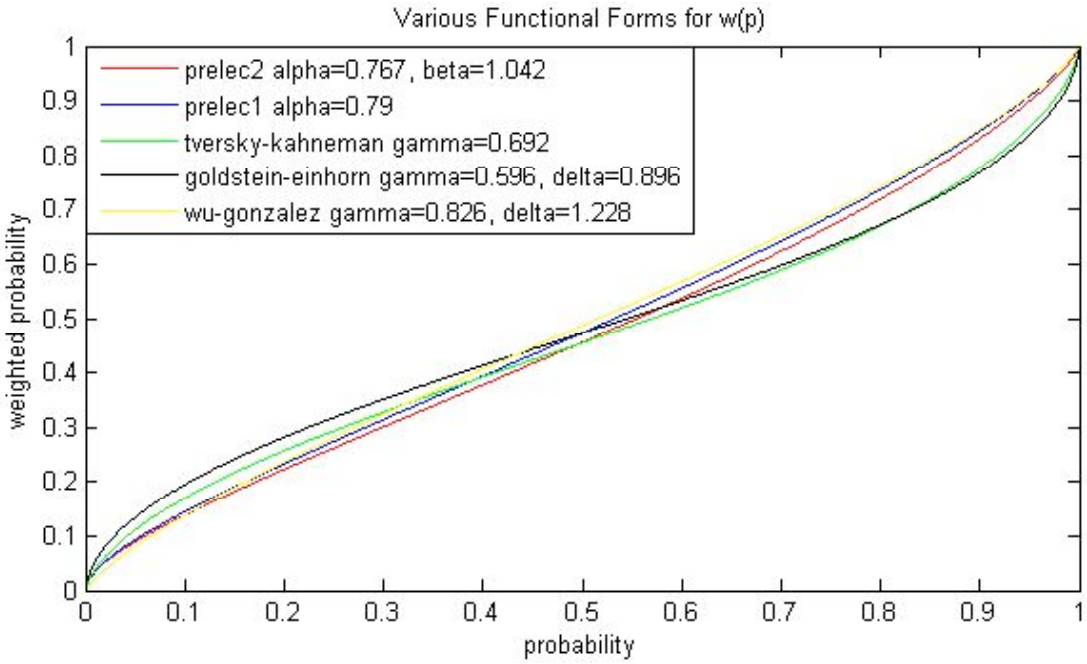


Figure 34: Probability-Weighted Probability plot of estimates from various studies using the functional forms used in those studies.

[10] Table 16, Table 17, and Figure 34 were provided by Chen Zhao, who provided excellent research assistant during her time as summer undergraduate research fellow at Caltech.

Functional Form	Study	Domain	Estimate
$w(p) = \frac{p^{\gamma}}{\left(p^{\gamma} + (1-p)^{\gamma}\right)^{-\gamma}}$	Tversky and Kahneman (1992)	Gains	$\gamma = 0.61$
		Losses	$\gamma = 0.69$
		Gains	$\gamma = 0.56$
	Camerer and Ho (1994)	Gains	$\gamma = 0.71$
	Wu and Gonzales (1996)	Gains	$\gamma = 0.60$
	Abdellaoui (2000)	Losses	$\gamma = 0.70$
		Health	$\gamma = 0.69$
	Bleichrodt, van Rijn, and Johannesson (1999)	Health	$\gamma = 0.71$
	Bleichrodt and Pinto (2000)	Gains	$\gamma = 0.96$
	Stott (2006)		
$w(p) = \frac{\delta p^{\gamma}}{\delta p^{\gamma} + (1-p)^{\gamma}}$	Wu and Gonzalez (1996)	Gains	$\delta = 0.84, \gamma = 0.68$
	Gonzalez and Wu (1999)	Gains	$\delta = 0.77, \gamma = 0.44$
	Tversky and Fox (1995)	Gains	$\delta = 0.77, \gamma = 0.69$
	Abdellaoui (2000)	Gains	$\delta = 0.65, \gamma = 0.60$
	Abdellaoui (2000)	Gains	$\delta = 0.84, \gamma = 0.64$
	Stott (2006)	Gains	$\delta = 1.40, \gamma = 0.96$
	Kilka and Weber (2001)	Stocks	$\delta = 1.036, \gamma = 0.443$
			$\delta = 0.928, \gamma = 0.312$
$w(p) = \exp(-(-\ln p)^{\alpha})$	Wu and Gonzalez (1996)	Gains	$\alpha = 0.74$
	Stott (2006)	Gains	$\alpha = 0.94$
	Tanaka, Camerer, Nguyen (forthcoming)	Gains	$\alpha = 0.72$ (students)
			$\alpha = 0.74$ (non-students)
$w(p) = \exp(-\beta(-\ln p)^{\alpha})$	Stott (2006)	Gains	$\beta = 1.00, \alpha = 1.00$
	Bleichrodt and Pinto (2000)	Health	$\beta = 1.084, \alpha = 0.534$
$w(p) = \frac{p^{\gamma}}{(p^{\gamma} + (1-p)^{\gamma})^{\ddot{a}}}$	Wu and Gonzalez (1996)	Gains	$\delta = 1.565, \gamma = 0.721$
	Stott (2006)	Gains	$\delta = 0.89, \gamma = 0.93$

Table 16: Summary of findings of past studies on probability weighting.

Study Name	Methodology	Parameters	Findings
Advances in Prospect Theory: Cumulative Representation of Uncertainty (Tversky and Kahneman, 1992)	Subjects indicated preference between one risky prospect and each of one set of seven sure outcomes that were shown logarithmically spaced between extremes of the risky prospect and each of a second set of seven sure outcomes shown linearly spaced between a value 25% higher than lowest amount accepted and value 25% lower than highest amount rejected. The certainty equivalent of the risky prospect was estimated as the average of the lowest accepted value and highest rejected value from the second set.	Outcomes: (0, 50), (0, -50), (0, 100), (0, -100), (0, 200), (0, -200), (0, 400), (0, -400), (50, 100), (-50, -100), (50, 150), (-50, -150), (100, 200), (-100, -200) Probabilities: 0.01, 0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95, 0.99	The value and weighting functions imply risk-aversion for gain and risk-seeking behavior for losses, both for prospects with moderate to high probability. The weighting function shows, for non-extreme outcomes, risk-seeking behavior for gains and risk aversion for losses. Estimated γ = 0.61 (gains) and γ = 0.69 (losses) for functional form of w(p) in current paper.
Violations of the Betweenness Axiom and Nonlinearity in Probability (Camerer and Ho, 1994)	Subjects chose from gambles represented both in compound form and reduced form. Gambles were grouped in gamble triples and each had a PL, PM, PH, the probabilities associated with the worst, middle and best outcomes.	(PL, PM, PH): (0.3, 0.4, 0.3), (0.4, 0.2, 0.4), (0.5, 0.0, 0.5), (0.4, 0.6, 0.0), (0.5, 0.4, 0.1), (0.6, 0.2, 0.2), (0.7, 0.0, 0.3), (0.66, 0.34, 0.00), (0.67, 0.32, 0.01), (0.83, 0.00, 0.17)	Estimated γ = 0.52 and γ = 0.56 for CPT form of w(p) introduced in Tversky and Kahneman (1992). Low probabilities are overweighted and high probabilities are underweighted. The crossover point where w(p)=p is around p=0.3.
Weighing Risk and Uncertainty (Tversky and Fox, 1995)	Certainty equivalents were inferred through series of subjects' choices between risky prospects and sure payments. Chance prospects concerned NBA playoffs, the Super Bowl, the future value of the Dow Jones index, and future temperatures in San Francisco and Beijing.	Probabilities varied between 0.05 and 0.95 in multiples of 0.05.	Estimated parameters for functional form from Lattimore, Baker and Witte (1992) are δ = 0.77, γ = 0.69 (chance), δ = 0.76, γ = 0.69 (Super Bowl), and δ = 0.72, γ = 0.76 (Dow Jones).
Probability Weighting and Utility Curvature in QALY-Based Decision Making (Bleichrodt and van Rijn, 1999)	Subjects faced gamble with certain option of living for 30 years with severe lower back pain and treatment option of (30 years in full health, p; immediate death , 1-p). Subjects indicated p for which they were indifferent between two options. Subjects also ranked seven health profiles describing years in full health, followed by years with severe lower back pain, followed by death.	Health Profiles: (Years in full health, Years with severe lower back pain) (0, 20), (18, 0), (16, 0), (14, 0), (12, 0), (8, 8), (6, 11)	Estimated γ = 0.69 for Tversky and Kahneman (1992) functional form. Quality-Adjusted Life-Years (QALYs) with probability weighting significantly more consistent with directly elicited rankings than QALYs without.
Curvature of the Probability Weighting Function (Wu and Gonzalez, 1996)	Eight pairs of gambles were arranged in "ladders" in which each step adds a common consequence to both gambles. Subject made choices for four rungs on each ladder.	Ladder 1: R1: (0.05, 240), (0.07, 200) R2: (0.05, 240), (0.10, 200), (0.17, 200) R3: (0.05, 240), (0.20, 200), (0.27, 200) R4: (0.05, 240), (0.30, 200), (0.37, 200) R5: (0.05, 240), (0.45, 200), (0.52, 200) R6: (0.05, 240), (0.60, 200), (0.67, 200) R7: (0.05, 240), (0.75, 200), (0.82, 200) R8: (0.05,	Estimated parameter for Tversky and Kahneman (1992) functional form is γ =0.71, for Prelec (1998) form is γ =0.74, for Goldstein and Einhorn (1987) form is δ = 0.84 and γ = 0.68, and form from current paper is δ = 1.565 and γ = 0.721. Weighting function is concave up to p≈0.40 and convex afterwards. A

		240), (0.90, 200), (0.97, 200) Ladder 2: R1: (0.05, 100), (0.10, 50) R2: (0.05, 100), (0.10, 50), (0.20, 50) R3: (0.05, 100), (0.20, 50), (0.30, 50) R4: (0.05, 100), (0.30, 50), (0.40, 50) R5: (0.05, 100), (0.45, 50), (0.55, 50) R6: (0.05, 100), (0.60, 50), (0.70, 50) R7: (0.05, 100), (0.75, 50), (0.85, 50) R8: (0.05, 100), (0.90, 50), (1.0, 50) Ladder 3: R1: (0.01, 300), (0.02, 150) R2: (0.01, 300), (0.10, 150), (0.12, 150) R3: (0.01, 300), (0.20, 150), (0.22, 150) R4: (0.01, 300), (0.30, 150), (0.32, 150) R5: (0.01, 300), (0.45, 150), (0.47, 150) R6: (0.01, 300), (0.60, 150), (0.62, 150) R7: (0.01, 300), (0.80, 150), (0.82, 150) R8: (0.01, 300), (0.98, 150), (1.0, 150) Ladder 4: R1: (0.03, 320), (0.05, 200) R2: (0.03, 320), (0.10, 200), (0.15, 200) R3: (0.03, 320), (0.20, 200), (0.25, 200) R4: (0.03, 320), (0.30, 200), (0.35, 200) R5: (0.03, 320), (0.45, 200), (0.50, 200) R6: (0.03, 320), (0.60, 200), (0.70, 200) R7: (0.03, 320), (0.85, 200), (0.90, 200) R8: (0.03, 320), (0.95, 200), (1.0, 200) Ladder 5: R1: (0.01, 500), (0.05, 100) R2: (0.01, 500), (0.10, 100), (0.15, 100) R3: (0.01, 500), (0.20, 100), (0.25, 100) R4: (0.01, 500), (0.30, 10), (0.35, 100) R5: (0.01, 500), (0.45, 100), (0.50, 100) R6: (0.01, 500), (0.65, 100), (0.70, 100) R7: (0.01, 500), (0.80, 100), (0.85, 100) R8: (0.01, 500), (0.95, 100), (1.0, 100)	weighting function that is strictly nonlinear with the boundaries outperforms a linear weighting function with discontinuities at 0 and 1.
A Parameter-Free Elicitation of the Probability Weighting Function in Medical Decision Analysis (Bleichrodt and Pinto,	Subjects were asked to choose a outcomes for which they were indifferent between two treatments that offered varying chances of surviving some x number of years.	Probabilities: 0.10, 0.25, 0.50, 0.75, 0.90	Estimated parameters for Tversky and Fox (1992) functional form is $\gamma = 0.674$, for Gonzalez and Wu (1999) form is $\gamma = 0.550$

2000)			and $\delta = 0.816$, for Prelec (1998) one parameter form is $\alpha = 0.533$, for Prelec (1998) two parameter form is $\alpha = 0.534$ and $\beta = 1.083$. There is significant evidence at aggregate and individual level for a nonlinear probability weighting function with crossover point between 0.25 and 0.50. Also, there is more evidence for elevation of w(p) than in previous studies.
On the Shape of the Probability Weighting Function (Gonzalez and Wu, 1999)	Subjects were presented with one gamble and a list of sure bets with differing outcomes. Subjects indicated preference for either gamble or sure thing for each sure bet in list.	Two outcome gambles: 25-0, 50-0, 75-0, 100-0, 150-0, 200-0, 400-0, 800-0, 50-25, 75-50, 100-50, 150-50, 150-100, 200-100, 200-150 Probabilities: 0.01, 0.05, 0.10, 0.25, 0.40, 0.50, 0.60, 0.75, 0.90, 0.95, 0.99	Estimates of parameters are $\delta = 0.77$ and $\gamma = 0.44$ for functional form in Goldstein and Einhorn (1987). At the individual level, the inverse S-shape of w(p) is regular, but there is much heterogeneity in terms of degree of curvature and elevation.
What Determines the Shape of the Probability Weighting Function Under Uncertainty? (Kilka and Weber, 2001)	Subjects were asked to choose between one risky asset (Deutsche Bank and Dai-Ichi Kangyo stocks) and a series of sure amounts and then to indicate the exact certainty equivalent. Lastly, subjects were asked to asses the probability of an event occuring.	Sure payments: from 10 DM to 110 DM in steps of 10 DM	Parameter estimates for the Gonzalez and Wu (1999) functional form is $\delta = 1.036$ and $\gamma = 0.443$ (Deutsche Bank) and $\delta = 0.928$ and $\gamma = 0.312$ (Dai-Ichi Kangyo).
Preferences, Poverty, and Politics: Experimental and Survey Data From Vietnam (Tanaka, Camerer, and Nguyen, forthcoming)	Subjects (Vietnamese villagers) were presented with a series of two options, the first of which stayed the same and the second of which increased in expected value.	Series 1: Option A: (0.3, 40000; 0.7, 10000) Option B: (0.1, x; 0.9, 5000) with x = 68000, 75000, 83000, 93000, 106000, 125000, 150000, 185000, 220000, 300000, 400000, 600000, 1000000, 1700000 Series 2: Option A: (0.9, 40000; 0.1, 30000) Option B: (0.7, x; 0.3, 50000) with x = 54000, 56000, 58000, 60000, 62000, 65000, 68000, 72000, 77000, 83000, 90000, 100000, 110000, 130000 Series 3: Option A: (0.5, x; 0.5, y) with x = 25000, 4000, 1000 and y = -4000, -8000 Option B: (0.5, 30000; 0.5, y) with y = -	Estimated parameter for one parameter Prelec (1998) functional form is $\alpha = 0.72$ for students and $\alpha = 0.74$ for non-students. Males have a more inflected weighting function with lower α. The parameter α also decreases with as the mean village income increases.

Cumulative Prospect Theory's Functional Menagerie (Stott, 2006)	Subjects were given a series of choices between two risky prospects.	21000, -16000, -14000, -11000 Outcomes: 0, 2500, 5000, 10000, 20000, 40000 Probabilities: 0.1 to 0.9 in increments of 0.1	Parameters estimates for the Goldstein-Einhorn (1987) functional form is $\delta = 1.40$ and $\gamma = 0.96$, Tversky and Kahneman (1992) functional form is $\gamma = 0.96$, Wu and Gonzalez (1999) is $\delta = 0.89$ and $\gamma = 0.93$, one parameter Prelec (1998) functional form is $\alpha = 0.94$, and two parameter Prelec (1998) functional form is $\beta = 1.00$, $\alpha = 1.00$. One parameter functions are preferable to two parameter versions (i.e. one parameter Prelec is better than two parameter Prelec).

Table 17: Summary of stimuli used and experimental and estimation procedure in past studies on probability weighting.

Proportion Judgments in Other Domain

Probabilities are often presented in experiments and thought of in life as proportions. It is therefore natural to ask to what extent results in probability weighting can be extended to proportion judgments. This pattern of bias in proportion and, more generally, ratio judgments has been found in a surprisingly wide variety of tasks and modality. Varey, Mellers, and Birnbaum (1990) used a task where subjects estimated the proportion of black and white dots in a field of black and white dots. Erlick (1964) showed participants a series of letters (*A*s and *C*s) at a rate of four letters per second and varied the proportion of *A*s. Nakajima (1987), in the auditory domain, asked participants to estimate the ratio of the duration of two silent time intervals demarcated by clicks. Figure 35 shows the judged proportion and bias as a function of true proportion in the three studies. All found that people consistently overestimated the small proportions and underestimated large ones (for review, see Hollands and Dyre 2000).

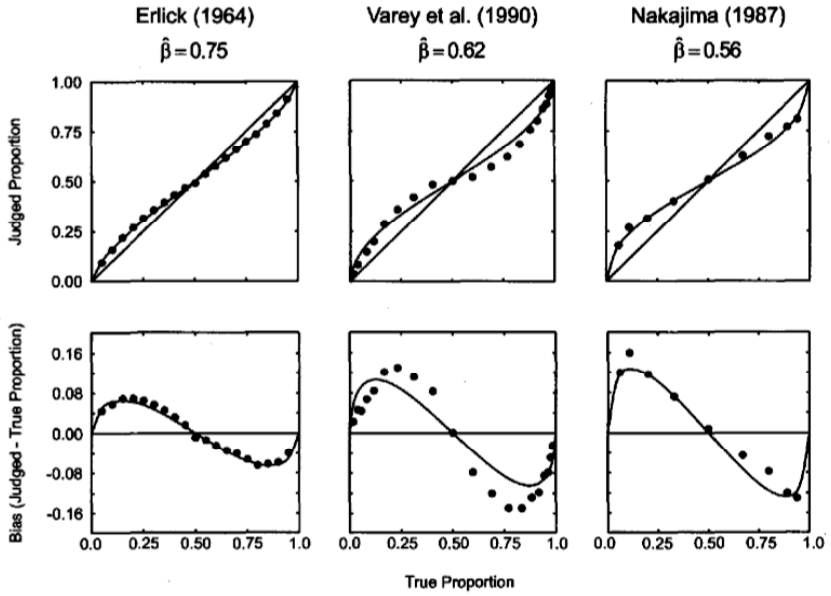


Figure 35: Judged proportion and bias as a function of true proportion for (A) letters over time (Erlick 1964), (B) mixed black and white dots (Varey et al. 1990), and (C) brief silent intervals (Nakajima 1987). Adapted from Hollands and Dyre, (2000).

Somewhat amazingly, there have been robust documentations of *multiple* inflection points—a phenomenon Hollands and Dyre calls "multiple-cycles". These patterns, which look like the probability weighting function repeated twice, appears in cases where participants are asked to judge proportion of pie charts or bars.

Derivation from the Power Law

Many of the works on the models of the psychophysics of proportions take as their starting point the well-known power law of psychophysics (often referred to as "Steven's law"): $\phi = \alpha \cdot x^{\beta}$ where ϕ is the perception of the stimulus quantity x, and α is a scale factor and β the exponent of the power function.

There is much evidence that the perception of numbers follows the power law at the behavioral level, and more recently, at the neural level as well, for both human and primates (Dehaene 2003; for review see Nieder and Miller 2003; Nieder 2005). These studies show that both cells in the prefrontal cortex and the intraparietal sulcus are differentially tuned to specific numbers, and that the distributions of the errors are lognormal. Their findings are also concordant with lesion studies showing that damage to the intraparietal sulcus results in deficits in numerical processing, a syndrome called *acalculia* (Dehaene and Cohen 1997).

Taking the power law as a starting point, and supposing that each quantity was perceived in such a way, the perception of the proportion π may be modeled as

$$\pi = \frac{\alpha p^{\beta}}{\alpha p^{\beta} + \alpha (1-p)^{\beta}} = \frac{1}{1 + \left[(1-p)/p\right]^{\beta}},$$

A different derivation of this functional form is obtained by using the ratio of the quantities π and $1 - \pi$ (Spence 1990). That is,

$$\frac{(1-\pi)}{\pi} = \frac{\alpha(1-p)^{\beta}}{ap^{\beta}} = \left((1-p)/p\right)^{\beta}$$

$$\frac{1}{\pi} - 1 = \left((1-p)/p\right)^{\beta}$$

$$\pi = \frac{1}{1+\left((1-p)/p\right)^{\beta}}.$$

A number of studies have used this parametric form (which is a constrained version of the Lattimore et al. 1992). Table 18 summarizes some of the estimates from this family of literature (which is quite diverse, from human-computer interaction, to engineering, to psychology). The results are surprisingly congruous with that of the probability weighting literature (see Table 16).

Study	Task	Estimate
Varey (1990)	Proportion of white and black dots	0.62
Erlich (1964)	Proportion of letters over time	0.75
Nakajima (1987)	Ratio of duration of time demarcated by sound	0.56
Spence (1990)	Proportion of three-dimensional bar graph	0.86
	Proportion of vertical line graphs	0.76
	Proportion of horizontal line graphs	0.87
	Proportion of pie chart	0.75
	Proportion of bar graph	0.76

Table 18: Estimates from studies of proportion judgment using the function $\pi = \frac{p^{\gamma}}{p^{\gamma} + (1-p)^{\gamma}}$.

As a further test of whether a common Power Law can be use to explain both the perception of proportion *and* magnitude, Hollands and Dyre (2000) correlated the estimates from various magnitude estimation and proportion judgment tasks. It was found that the parameters are correlated at $\rho = 0.91$ (Figure 36).

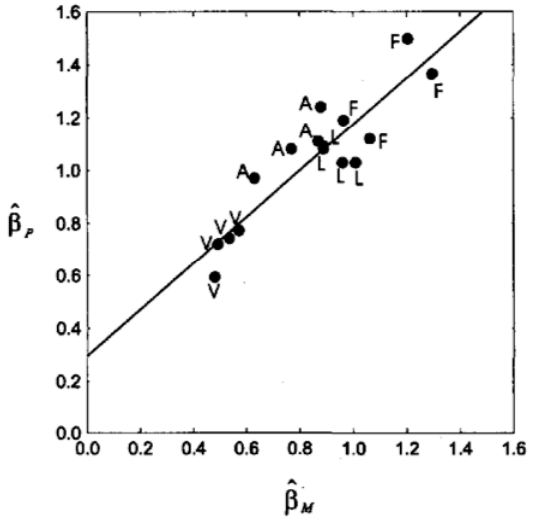


Figure 36: Correlation of estimate of exponent from power function in magnitude estimation (x-axis) and in proportion judgments (y-axis). Key: V = volume; A = area; L = length; F = finger span. Adapted from Hollands and Dyre, (2000).

Interestingly, the variants of the above functional form were proposed as early as 1978 in the probability weighting literature (Karmarkar 1978; Lattimore et al. 1992; Gonzalez and Wu 1999), albeit not explicitly derived from the Power Law.

It is surprising that given that the literature on the psychophysics of proportion predates that of probability weighting, the latter does not appear to be aware of the former (the reverse, however, appears not to be true, as Hollands and Dyre cites Tversky and Kahneman, 1992). Indeed, it was uncited in the prominent studies of probability weighting (e.g. Lattimore et al. 1992; Tversky and Kahneman 1992; Wu and Gonzalez 1996). The parallels, as discussed earlier, are striking.

Animal Literature

The study of nonlinear weighting of probabilities has been, to my knowledge, exclusively limited to humans. The animal decision literature has, in general, ignored such second-order effects such as nonlinearity of the probabilities. However, some clues may be found in their data. Sugrue et al. (2005) provides one such example. They used a matching-pennies game and found that monkeys were able to behave much as the game theoretic solution would have predicted. Notably, however, and not mentioned by the authors, the monkey's matching behavior looks remarkably similar to the human data.

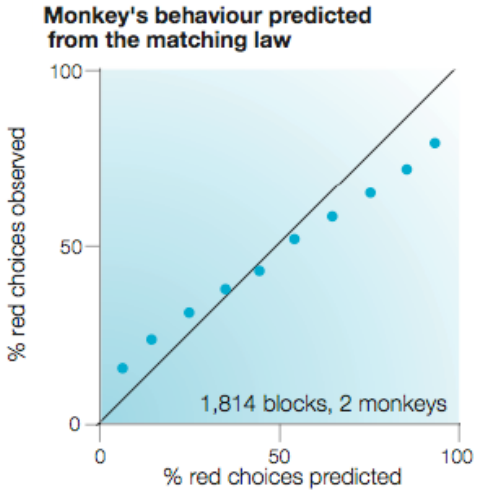


Figure 37: Average monkey behavior in a probability-matching task. Standard error bars are smaller than data points and therefore not visible. Adapted from Sugrue et al. (2005).

This data is notable on two accounts. First, it suggests that probability weighting generalizes across species. Second, and more importantly, this data suggests that the nonlinearity generalizes to both experienced and described probabilities. This is because experimental work with monkeys cannot use abstract, cognitive depictions of probabilities. Instead, monkeys learn probabilities through many hundreds, or thousands, of trials. This is as opposed to some of the findings in the human literature that experienced probabilities are weighted differently than those of described probabilities (Hertwig et al. 2004), and suggests that, given adequate sampling across the probabilities, experienced and described probabilities are weighted similarly (i.e., the inverse S-shape).

Neural Hypotheses

From the above, it appears that there are two biases that take place with the weighting of probabilities in decisions under risk. The first is that the brain is the perception of probabilities. This appears to produce the inverse S-shape that has been found so consistently in the literature on probability weighting *and* proportion judgment. The second is the elevation of the inverse S-shape curve, which is *not* consistent with the literature on proportion judgment.

In the probability weighting literature, the location of the point at which the weighting function crosses the diagonal is between 0.3 and 0.4. In the proportion perception literature

this is around 0.5. Therefore, whereas the curve observed in the latter is symmetric (i.e., the deviation from the diagonal is approximately 0), it is asymmetric in the former. A psychological interpretation of these parameters, given by Gonzalez and Wu (1999), states that the curvature of the weighting function is a product of the *perception* probabilities, whereas the elevation is one of the evaluation of the attractiveness of the outcome.

This interpretation is neurally plausible, as the perception of proportions and the integration of probabilities and rewards are known to involve different, but overlapping, sets of regions in the brain. Many studies focusing on numerical processing finds that the intraparietal lobule is involved consistently (Dehaene 2003; Hubbard et al. 2005). The processing of expected reward and reward learning, in contrast, is generally thought to be critically subserved by areas that are projected to by the midbrain dopaminergic neurons—in areas such as the dorsal and ventral striatum.

One possibility is that in the course of combining probabilities and rewards, the probabilities undergoes another transformation, i.e., $w(p) = \left(\frac{p^{\gamma}}{p^{\gamma} + (1-p)^{\gamma}} \right)^{\delta}$. This function can produce much of the same behavior as the three functions that were reviewed earlier (see Figure 38).

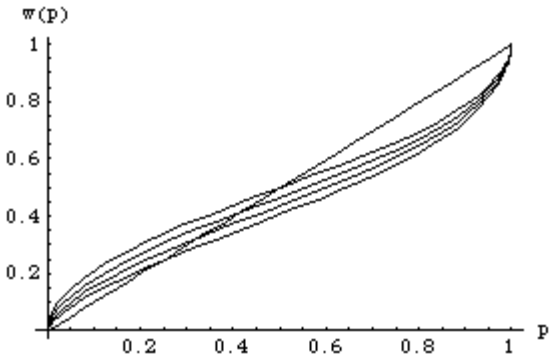


Figure 38: Plot of $w(p) = \left(\frac{p^{\gamma}}{p^{\gamma} + (1-p)^{\gamma}} \right)^{\delta}$ with γ constrained to 0.6 and δ varying between 1 and 1.3.

Another possibility is to derive the Lattimore function from the Power Law by assuming that either the scaling factor α, or the exponent β, is variable in $\phi = \alpha \cdot x^{\beta}$. With a variable scaling factor, the weight $\delta = \alpha_1/\alpha_2$; this quantity represents the ratio of the scaling factors between the alternatives. With a variable β, the parameter $\delta = 1/(1-p)^{\gamma_2 - \gamma_1}$. Because of this equivalence, it is not possible to discriminate between these processes unless one can estimate the underlying power function.

This view puts the neural correlate of the empirical probability weighting function at the expected reward regions, rather than the perception regions. The prime candidates for such a function are the dopaminergic areas that are known to be involved in reward anticipation and learning.

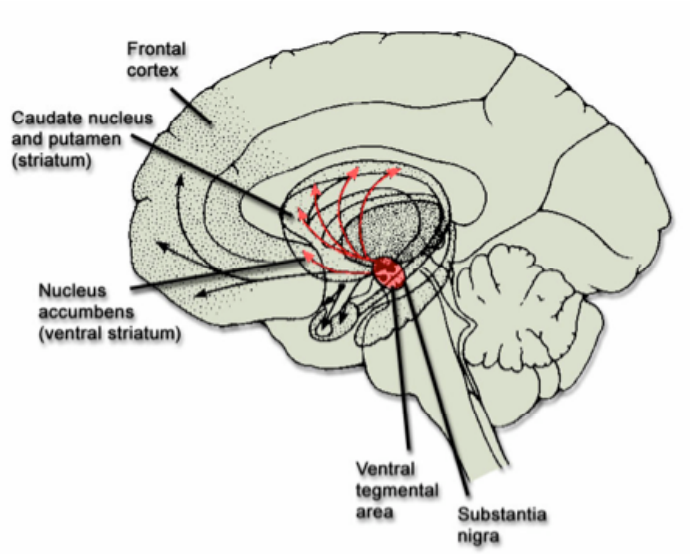


Figure 39: Dopamine and its projections in the brain. (Adapted from Kandel et al., 1991.)

This was hypothesized and demonstrated in a series of beautiful and ground-breaking papers by Wolfram Schultz and colleagues (Schultz et al. 1997; Schultz et al. 2000; Fiorillo et al. 2003; Tobler et al. 2005). They were able to show that midbrain dopaminergic

neurons encoded both expected reward (reward prediction) as well as the reward prediction error (see Figure 40 and caption for details).

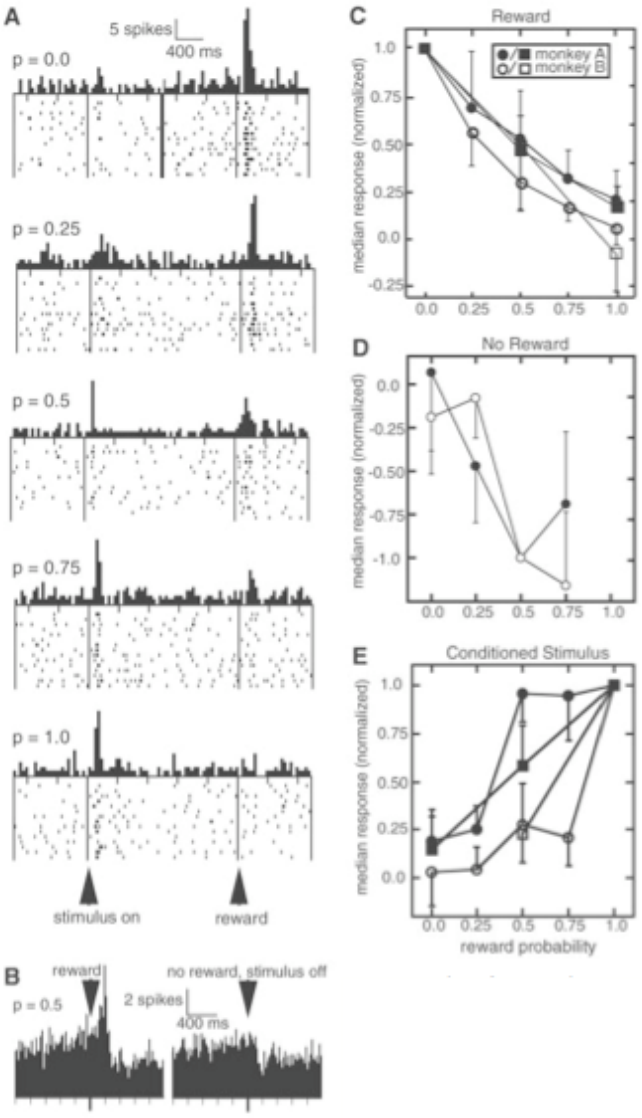


Figure 40: Response of midbrain dopaminergic neurons to juice delivered in trials with varying probabilities of delivery. (A) Raster and histogram of activity of a single cell over various levels of probabilities. Only data from rewarded trials were included. (B) Population histogram of activity responding to rewarded (left) and unrewarded (right) trials at p=0.5 of juice delivery. (C) Median response to juice delivery as a function of probability of delivery. Error bars are standard error of the mean (SEM). Circles and squares are from data from different experiments. (D) Same as (C), except only activity from unrewarded trials are plotted. (E) Same as (C) and (D), except

data plotted are those from the conditioned stimulus indicating the probability of juice delivery. Adapted from Fiorillo et al. (2003).

This result has been replicated in humans with fMRI, which showed additional activation in the putamen and the caudate, both part of the dorsal striatum (Preuschoff et al. 2006).

Other candidate regions include the lateral intraparietal cortex, which has been shown to respond to expected reward (Platt and Glimcher 1999) as well as encoding and comparison of numerosity (Nieder and Miller 2003).

Experiment Design

Subjects

Subjects were recruited from the online bulletin board, the Los Angeles portion of Craig's List (http://losangeles.craigslist.org/) under the categories "community: volunteers", "community: general", "gigs: event", and "gigs: labor".

The experiment consisted of 22 subjects (11 female). Mean (standard deviation) age is 29.6 (7.5). Informed consent was given through a consent form approved by the Internal Review Board at Caltech. Subjects were given a $5 show-up fee.

Behavioral Task

Subjects were given written instructions, which were also read out loud to them by the experimenter. They were then given a quiz to assess comprehension. Immediately afterwards, each subject completed six practice rounds with gambles and probabilities that were not used for the actual experiment.

The experiment consisted of 120 short, self-paced rounds. In each round, subjects had the task of choosing between two simple gambles (receive x dollars with probability p and zero dollars with probability 1 - p).

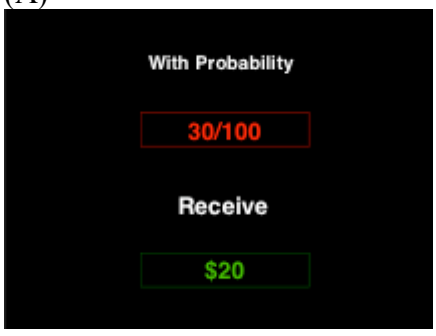
The first screen is a fixation screen that randomly lasts 6, 8, or 10 seconds. The second screen is always shown for 4 seconds. This shows the probability (p_1) and outcome value (x_1) for the first gamble.

On one-tenth of the rounds, the subjects were asked an attention check question on the next screen: whether p_1 was greater than or less than 0.4. This is to ensure that the subjects paid attention to the first screen. The fixation screens between the second and third screens are randomly either 4 or 6 seconds.

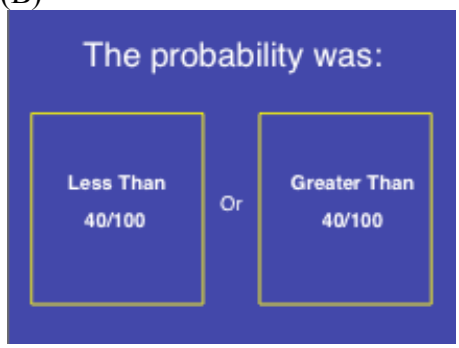
Then, on all the trials, the last screen asks them to choose between the first gamble and a second gamble. The second and third screens, that require responses, are self-paced. The fixation screens between the first and second screens and between the second and third screens are randomly either 4 or 6 seconds. After subjects enter their response, the box around the option that they choose turns red to confirm their choice.

The next screen displayed both gambles with the probabilities (p_1 and p_2) and outcome values (x_1 and x_2) and prompted the subjects to choose a gamble. The relative position (left and right) of the two gambles was randomized to counterbalance for order effects. Subjects entered responses via a button box with two buttons (left button for left option on screen; right button for right option). Subjects saw either two or three screens on each round. Figure 41 shows the three screens that are presented to the subjects.

(A)

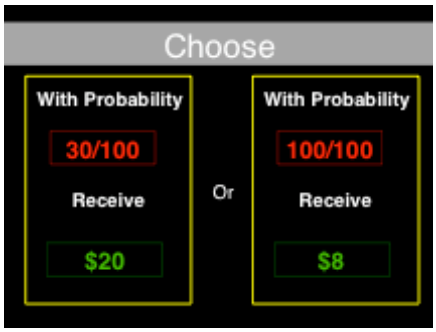


(B)



(C)

Figure 41: The three screens from the experiment. (A) The first screen that the



subjects saw with information about the first gamble; (B) the attention-check question screen that was displayed on one tenth of the trials; (C) the choice screen that displayed information about both gambles.

The entire experiment has 120 trials divided into 3 blocks. At the end, one round was randomly chosen using a deck of cards numbered 1 through 120 to play out according to the subject's decision during that round. Gambles were played out using a deck of 100 cards numbered 1 through 100. Winning cards were designated by the probability for the gamble being played out such that probability = p/100 means that cards 1 through p were winning cards and cards p+1 through 100 were losing cards. Subjects were paid $20 for participating, $5 if they answer all the questions concerning whether the probability was greater or less than 40/100 correctly, and whatever they win from the gamble that is played.

Choice of (p, x)

To separate probability from outcome value and expected value, we fixed x_1 at four different levels, $x_1 \in \{10,25,50,100\}$, and sampled at six different probability levels, $p_1 \in \{0.01,0.1,0.3,0.5,0.8,0.95\}$, fives times each. The values for gamble 2, (p_2, x_2), were chosen such that the expected values for the two gambles were close and the expected weighted utility calculated using the Prelec function with α = 0.7 and a power utility function with $\rho = 0.5$ for the two gambles were such that for a particular gamble 1, subjects were expected to divide their choices between the two gambles roughly evenly.

At the end of the experiment, subjects were paid a base rate plus earnings from playing out the gamble that they chose on one randomly selected round and a reward contingent on their answering all the questions correctly. Average earning from the gambles were $19.70 plus the show-up fee of $5, not including the $5 bonus for remembering whether the probability of the gambles were greater or less than 40/100.

fMRI Data Collection

Scans were acquired using the 3 Tesla Siemens Trio scanner at Caltech's Broad Imaging Center. Anatomical images (high resolution, T1-weighted) were acquired first. Functional (T2-weighted) images were then acquired using the following parameters: TR= 2000ms, TE = 40ms, slice thickness = 4mm, 32 slices. Horizontal slices were acquired approximately 30° clockwise of the anterior-posterior commissure (AC-PC) axis to minimize signal dropout of the orbitofrontal cortex (John O'Doherty, personal communication). The total duration of the experiment varied, as each round was self-paced.

Behavioral Results

Estimation Procedure. A stochastic choice model was used to estimate the probability weighting function. For more details about the implementation of the procedure, see page 38.

Similar to a majority of the studies on this topic, we start by assuming the functional forms for the utility and probability weighting functions. We assume that the utility function is a power function $u(x;\rho) = x^{\rho}$ and the probability weighting function the Prelec's one parameter weighting function $\pi(p;\alpha) = \exp\{-\ln(-\ln p)^{\alpha}\}$. Expected utilities are assumed to combine decision weights and utilities linearly, $U(x,x',p;\rho,\alpha) = \pi(p,\alpha)u(x,\rho) + (1-\pi(p,\alpha))u(x',\rho)$.

Behavioral Estimates. Table 19 presents the parameter estimates for the stochastic choice model. These parameters are generally similar to those found in the literature reviewed in the earlier sections, with one notable difference. Five of the subjects follow a clear probability maximizing strategy. That is, they select the gamble that had the higher probability, regardless of the expected value of the gambles. Because of this pattern of behavior, it was impossible to characterize their behavior in our stochastic choice model with the given range of probability and stakes used in our experiment.

Subject	α	ρ	λ
23	0.833	0.421	5.025
24*	-	-	-
26	0.727	0.598	2.404
27	0.603	0.445	5.979
28	0.627	0.441	1.667
29*	-	-	-
30	0.613	0.486	2.708
31	0.759	0.469	6.491
32	0.693	0.73	1.016
33	1.62	0.722	0.875
34*	-	-	-
35*	-	-	-
35*	-	-	-
37	0.772	0.905	0.716
38	0.557	0.727	0.553
39	0.444	0.397	5.368
41	0.413	0.728	0.547
42	0.888	0.616	1.424
43	0.686	0.49	4.979
44	1.345	0.384	1.743
45	0.751	0.66	1.319
Mean	0.771	0.576	2.676

Table 19: Stochastic choice model estimates of probability weighting function (Prelec's one parameter model) and utility function (power function). The five probability-maximizers are denoted by *.

No gender differences were observed behaviorally in either probability weighting parameter (male $\alpha = 0.750$, female $\alpha = 0.695$, t = 0.47, $p < 0.64$ two-tailed) or risk aversion (male $\rho = 0.539$, female $\rho = 0.513$, t = 0.24, $p < 0.81$ two-tailed).

fMRI Results

The fMRI data analysis techniques used are similar to those discussed in Chapter 2; that is, standard random effects models are used.

Log Encoding Model

This model explored the possibility that the brain does a log expansion of the expected value. That is, $U(\pi, x) = \log(\pi) + \log(x)$. This has the virtue of being computationally more tractable for the brain (Dayan and Abbott 2001).

Furthermore, if the decision weight π is the perception of proportions such as those reviewed on page 89, this would be equivalent to a power transformation of the proportion judgment function discussed in that section.

We estimated a model, therefore, by including both the log(p) and log(x) parameters. The event was defined as the entire epoch of the first presentation of gamble 1, which is 4 seconds. Other nuisance parameters included regressors for the other phases of the trial, including the choice screen and, when present, the question screens of whether probabilities are greater than 0.4.

Log(*p*)

Regions significantly activated by the logarithm of the probabilities include some well-known areas involved in reward anticipation, including the putamen (Figure 42) and the midbrain. The caudate was also activated, but at less stringent *p*-values. See Table 20 for a comprehensive list of the regions.

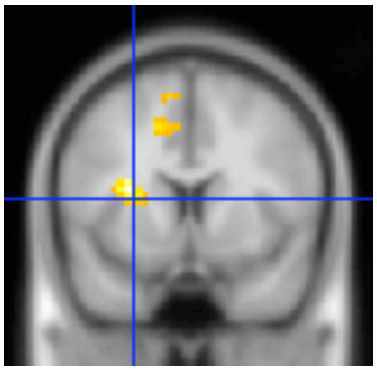


Figure 42: Putamen (y = 9) activity significantly correlated with log(p), $p < 0.001$, $k > 10$.

cluster level			voxel level									
p_{cor}	*k*	p_{unc}	p_{FWE}	p_{FDR}	*T*	*Z*	p_{unc}	*X*	*Y*	*Z*	*Side*	*Region*
0.002	139	0	0.012	0.009	6.96	4.96	0	-15	54	6	L	medial frontal gyrus
			0.758	0.056	4.3	3.6	0	-9	54	21	L	
			0.849	0.056	4.13	3.49	0	9	48	18	R	

0.078	55	0.016	0.023	0.009	6.6	4.81	0	-30	9	15	L	putamen
0.44	21	0.113	0.24	0.033	5.22	4.13	0	-24	-93	-6	L	occipital
0.039	70	0.008	0.38	0.038	4.91	3.96	0	-48	-24	-15	L	middle temporal gyrus
			0.654	0.056	4.46	3.7	0	-57	-21	-6	L	
			0.89	0.056	4.04	3.44	0	-63	-24	-15	L	
0.418	22	0.105	0.549	0.056	4.63	3.8	0	-48	-39	0	L	middle temporal gyrus
			0.964	0.056	3.8	3.27	0.001	-63	-36	0	L	
0.165	40	0.035	0.614	0.056	4.52	3.74	0	-3	-15	-15	L	midbrain
			0.79	0.056	4.24	3.56	0	-3	-24	-15	L	
			0.954	0.056	3.84	3.31	0	-3	-24	-27	L	
0.007	111	0.001	0.649	0.056	4.47	3.71	0	-12	-63	-24	L	cerebellum
			0.838	0.056	4.15	3.51	0	0	-66	-15	R	
			0.942	0.056	3.89	3.34	0	3	-57	-9	R	
0.134	44	0.028	0.741	0.056	4.33	3.62	0	-15	9	45	L	cingulate
			0.913	0.056	3.98	3.4	0	-9	9	60	L	
			0.938	0.056	3.9	3.35	0	-18	18	42	L	
0.712	11	0.242	0.805	0.056	4.22	3.55	0	0	36	-9	R	orbitofrontal cortex
0.463	20	0.121	0.854	0.056	4.12	3.49	0	-36	-84	-9	L	occipital
0.712	11	0.242	0.921	0.056	3.96	3.38	0	6	24	-6	R	anterior cingulate
0.712	11	0.242	0.931	0.056	3.93	3.36	0	39	-63	-15	R	subgyral

Table 20: Regions significantly correlated with log(p), $p < 0.001$, $k > 10$.

Log(x)

A much smaller set of regions was significantly activated by the logarithm of the gamble amounts. This is, perhaps, due to the fewer elements in the stakes set that we used (4 compared to 6, for probabilities. Nevertheless, it is clear that expected reward regions were significantly activated by the logarithm of the gamble stakes. This includes the caudate (Figure 43) and the putamen. See Table 21 for a comprehensive list of the regions.

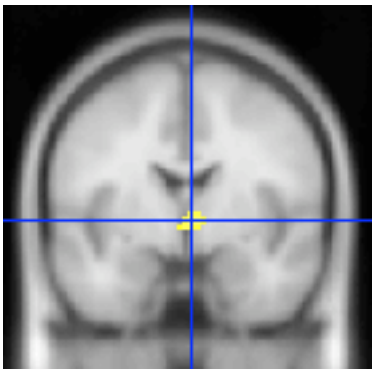


Figure 43: Caudate (y = -3) activity significantly correlated with log(x), $p < 0.001$, $k > 10$.

cluster level			voxel level									
p_{cor}	k	p_{unc}	p_{FWE}	p_{FDR}	T	Z	p_{unc}	X	Y	Z	*Side*	*Region*
0.288	36	0.101	0.264	0.1	4.92	3.97	0	33	3	33	R	putamen/subgyral
0.568	17	0.248	0.67	0.1	4.19	3.53	0	-54	-45	0	L	middle temporal gyrus

0.149	55	0.048	0.736	0.1	4.09	3.47	0	-15	21	60	L	superior frontal gyrus
			0.835	0.1	3.91	3.35	0	-21	12	63	L	premotor cortex
0.413	26	0.157	0.831	0.1	3.92	3.36	0	3	-3	0	R	caudate nucleus

Table 21: Regions significantly correlated with log(x), $p < 0.001$, $k > 10$.

Log(p)+Log(x)

A log-encoding model has the advantage that the expected value is the sum of the logarithms of the parts. Therefore, we look for the areas that are significantly activated by $\log(p) + \log(x)$. This set of regions includes the caudate nucleus and the midbrain, as well as the putamen (Figure 44). Table 22 presents the table of all activations at $p < 0.001$ and cluster size $k > 10$.

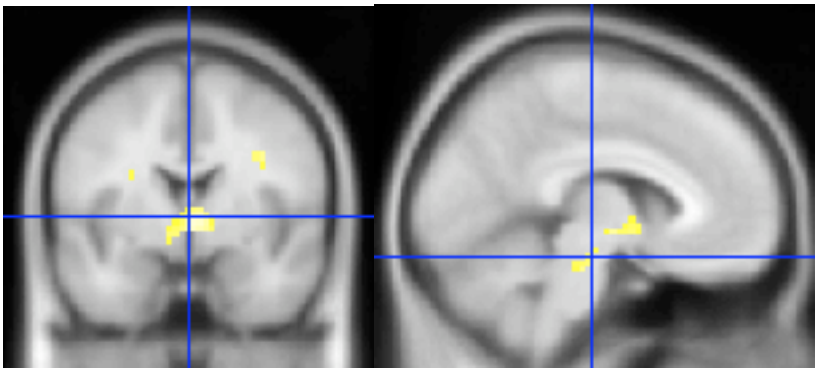


Figure 44: Regions significant correlated with log(p) + log(x), (Left) Caudate and (Right) midbrain. Both significant at $p < 0.001$, cluster size $k > 10$.

cluster level			voxel level									
p_{cor}	k	p_{unc}	p_{FWE}	p_{FDR}	T	Z	p_{unc}	X	Y	Z	*Side*	*Region*
0.149	55	0.048	0.317	0.053	4.8	3.9	0	-60	-21	-21	L	inferior temporal gyrus
			0.499	0.053	4.46	3.7	0	-57	-24	-12		
0.002	208	0.001	0.342	0.053	4.74	3.87	0	-18	18	60	L	middle frontal gyrus
			0.507	0.053	4.45	3.69	0	-42	21	45		
			0.81	0.053	3.96	3.38	0	-33	15	51		
0.3	35	0.106	0.427	0.053	4.58	3.77	0	-51	-45	0	L	middle temporal gyrus
0.048	90	0.015	0.544	0.053	4.39	3.66	0	3	0	0	M	caudate nucleus
			0.764	0.053	4.04	3.43	0	-6	-9	-6	R	
			0.868	0.053	3.84	3.3	0	-9	0	-6		
0.346	31	0.126	0.549	0.053	4.38	3.65	0	30	9	30	R	putamen/subgyral
0.203	46	0.067	0.551	0.053	4.38	3.65	0	-9	-18	-18	M	midbrain
			0.607	0.053	4.29	3.6	0	15	-30	-18		
			0.615	0.053	4.28	3.59	0	0	-24	-24		
0.512	20	0.213	0.741	0.053	4.08	3.46	0	-27	18	24	L	putamen/subgyral
			0.845	0.053	3.89	3.34	0	-27	3	21		
0.652	13	0.313	0.75	0.053	4.06	3.45	0	15	15	21		

0.674	12	0.333	0.808	0.053	3.96	3.38	0	-42	-78	3	L	middle occipital lobe
0.63	14	0.295	0.811	0.053	3.96	3.38	0	-30	-27	36	L	subgyral

Table 22: Regions significantly correlated with log(p) + log(x), $p < 0.001$, $k > 10$.

Power Utility Model

It is difficult to reject the log-encoding model and the power-utility model because the two make similar predictions about the neural activity. There is, therefore, little power to distinguish between these two models. Yet because they make such different predictions about the underlying neural mechanism, I include the results here.

As opposed to the previous log-encoding model, this model uses the utility functions estimated from the stochastic choice models. Because 5 of the subjects were probability maximizers, their utility function could not be estimated. They were, therefore, excluded from the analysis.

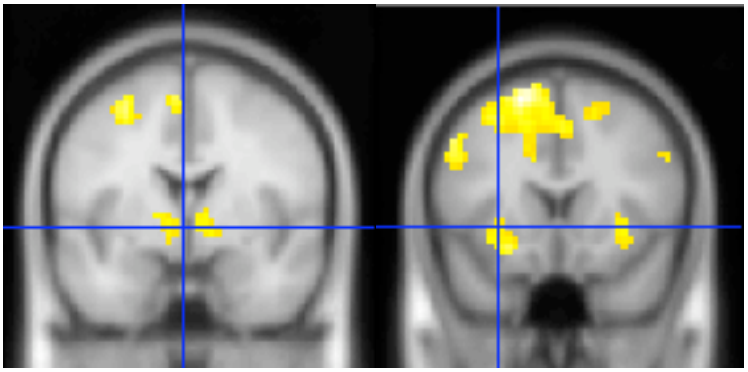


Figure 45: Regions significant correlated with $U(p,x;\rho) = p \bullet x^{\rho}$, (Left) Caudate and (Right) putamen. ($n = 17$, $p < 0.0005$, cluster size $k > 10$).

cluster level			voxel level									
p_{cor}	k	p_{unc}	p_{FWE}	p_{FDR}	T	Z	p_{unc}	X	Y	Z	*Side*	*Region*
0	911	0	0.029	0.01	6.45	4.64	0	-18	18	60	L	middle frontal gyrus
			0.082	0.01	5.79	4.34	0	-12	39	51	L	
			0.17	0.01	5.33	4.12	0	-18	42	42	L	
0	210	0	0.037	0.01	6.29	4.57	0	-60	-21	-21	L	inferior temporal gyrus
			0.212	0.01	5.18	4.04	0	-51	-24	15	L	
			0.454	0.01	4.62	3.74	0	-51	-30	-6	L	
0.025	78	0.013	0.167	0.01	5.34	4.12	0	-51	33	-3	L	lateral orbitofrontal cortex
			0.208	0.01	5.19	4.05	0	-42	45	-3	L	

			0.784	0.01	4.06	3.4	0	-39	54	6	L	
0.008	112	0.004	0.226	0.01	5.14	4.02	0	-9	-15	-12	L	midbrain/ventral striatum
			0.288	0.01	4.97	3.93	0	3	-21	15	R	
			0.541	0.01	4.47	3.65	0	-6	3	-3	L	caudate
0.064	54	0.032	0.289	0.01	4.97	3.93	0	-24	12	-12	L	Putamen
0.041	65	0.021	0.303	0.01	4.93	3.91	0	18	39	51	R	middle frontal gyrus
			0.434	0.01	4.66	3.76	0	15	21	51	R	
			0.634	0.01	4.32	3.56	0	18	9	57	R	
0.101	43	0.053	0.314	0.01	4.91	3.9	0	-27	54	3	L	superior frontal gyrus
			0.376	0.01	4.77	3.82	0	-15	57	6	L	
			0.528	0.01	4.49	3.66	0	-18	57	27	L	
0.144	35	0.077	0.369	0.01	4.79	3.83	0	-21	-93	-3	L	occipital cortex
			0.565	0.01	4.43	3.63	0	-33	-84	-9	L	
			0.615	0.01	4.35	3.58	0	-39	-78	-9	L	
0.292	20	0.171	0.381	0.01	4.76	3.82	0	6	24	-6	R	anterior cingulate
			0.7	0.01	4.21	3.49	0	6	21	-18	R	
0.031	72	0.016	0.387	0.01	4.75	3.81	0	51	24	21	R	brodman area 45
			0.431	0.01	4.67	3.76	0	51	18	33	R	
0.323	18	0.193	0.585	0.01	4.4	3.61	0	12	-51	21	R	posterior cingulate
0.151	34	0.081	0.588	0.01	4.39	3.61	0	9	0	-3	R	caudate
0.138	36	0.073	0.608	0.01	4.36	3.59	0	33	15	-3	R	putamen
0.461	11	0.305	0.671	0.01	4.26	3.52	0	-51	6	21	L	inferior frontal gyrus

Table 23: Regions significantly correlated with expected utility. Utility functions are calculated using power function and parameters estimated from individual subjects ($n = 17$, $p < 0.0005$, cluster size $k > 10$).

Probability Weighting Function

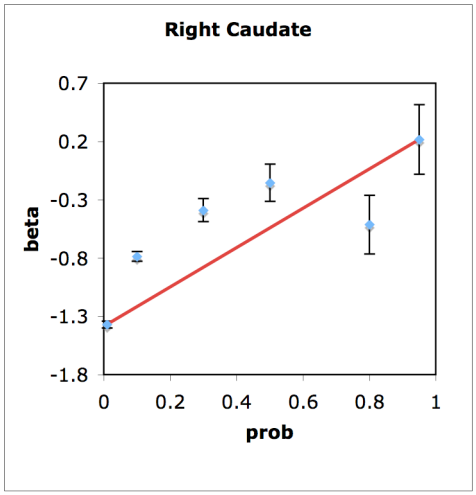
To search for the probability weighting function, a two-step procedure was used. First, the candidate regions were selected based on the results from the previous analyses and prior hypothesis. The caudate and the midbrain are distinguished as regions that clearly respond to both probabilities and rewards. This is clearly consistent with the idea that these regions encode probabilities and rewards.

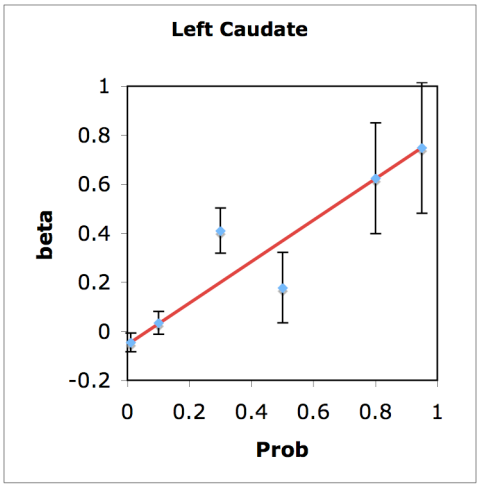
Regions of interests were selected at $p < 0.001$ uncorrected and cluster size $k > 10$. I selected the voxel with the highest activation in those clusters. I then specified a model where each individual gamble 1 was its own regressor. Beta values are defined for each probability level, using a stratified approach. That is, beta values are estimated for each stake that is involved in the experiment ($x = \{10, 25, 50, 100\}$). We therefore make the assumption that decision weighted and utilities are combined linearly. This is a common

assumption in non-expected utility theory, and in expected utility theory, decision weights are assumed to be the probabilities themselves.

This is because SPM uses a summary-statistics approach and does not allow the variance between the subjects to vary. In this case, the deviations from the SPM assumptions can be large because we're after quite a subtle nonlinearity. The approach used here allows me to estimate a true, mixed-effects model that takes into account the multiple layers of nested groupings—both sessions and subjects.

The beta for the brain area at each trial was then extracted and exported to *R*, an external statistical package, and estimated using the mixed effects package *nlme* (Pinheiro and Bates 2000). A nonparametric model was used to estimate the probability weighting function, where each probability level was given its own dummy variable. Figure 46 shows the beta-probability plot from these estimates in the caudate nucleus.





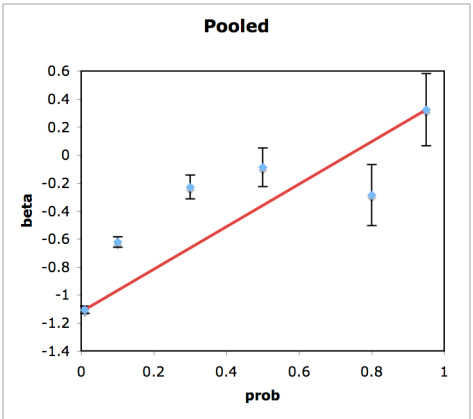


Figure 46: Beta-probability plot in the caudate nucleus.

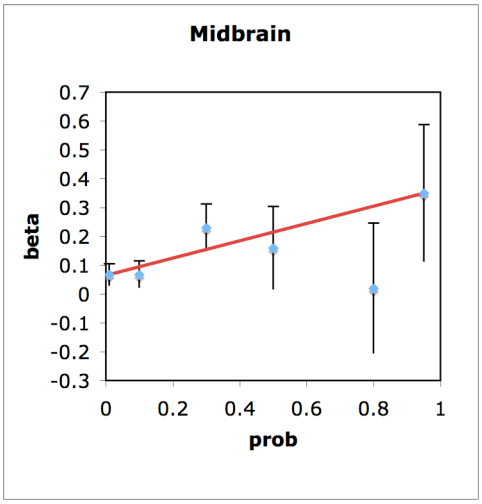


Figure 47: Beta-probability plot in the midbrain.

Laterality differences are not often noted in this literature. One possible reason for this is that much of the evidence on functions of the caudate and striatum come from animal work, especiaslly rats, but monkeys as well. The type of stimuli they are able to use, as a result, are restricted compared to those used in work in the decision-making literature, which almost exclusively use humans.

One example of laterality differences in the striatum (and caudate) specifically, is from the ambiguity study in Chapter 2. There we find that the activity in the left caudate is more highly correlated with expected value of the choices in the knowledge treatment, whereas the right caudate is more highly correlated with expected value in the card-deck treatment.

This suggests that the laterality difference that we see in Figure 44 may be due to numerical processing of the right hemisphere.

Using Difference Method

Because the subtly of the nonlinearity involved, another method was used to explore the nonlinear relationship. This method involved looking for areas that are significantly correlated with the residual—that is, the difference between the decision weights and the probabilities.

Denote the true weighting function for subject *i* to be $\pi(p,\alpha_i)$, and this difference measure $\Delta(p,\alpha_i) = \pi(p,\alpha_i) - p$. A brain region that represents $\pi(p,\alpha_i)$ will be significantly correlated ith both $\Delta(p,\alpha_i)$ *and p*. The regression equation therefore i

$$\pi(p,\alpha_i) = a + b_1 \cdot p + b_2 \cdot \Delta(p,\alpha_i) + e.$$

Method I: Maximum Likelihood $\hat{\alpha}_i$

Because we do not observe the true α_i, we do not know the true Δ_i. One solution is to straightforwardly replace the true α_i with the maximum likelihood $\hat{\alpha}_i$ derived from the behavioral data.

This method uses $\Delta(p,\hat{\alpha}_i) = \pi(p,\hat{\alpha}_i) - p$, where we calculate the difference using the best fit Prelec function for each subject *i*. This has the advantage of being intuitive and straightforward.

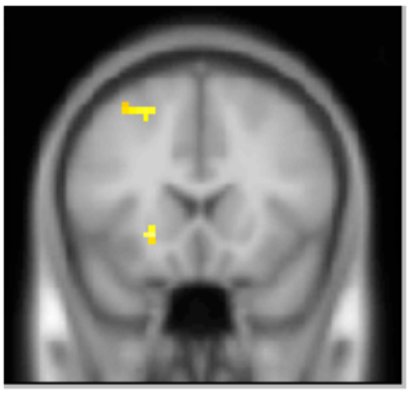


Figure 48: Regions significantly correlated with $\Delta(p,\hat{\alpha}_i) = \pi(p,\hat{\alpha}_i) - p$ include putamen (y = 18) at $p < 0.005$, $k > 5$.

k	p_{FWE}	p_{FDR}	*T*	p_{unc}	*X*	*Y*	*Z*	*L/R*	*Region*
39	0	0.018	15.35	0	-18	12	48	L	Superior Frontal Gyrus
			12.73	0	-12	15	60	L	
			12.28	0	-30	15	51	L	
4	0	0.018	13.64	0	-57	-21	-21	L	Inferior Temporal Gyrus
6	0	0.018	12.22	0	-21	18	-6	L	Putamen
6	0	0.018	11.77	0	-33	3	51	L	Brodman Area 6
12	0	0.018	11.42	0	3	45	27	R	Brodman Area 9

Table 24: Regions significantly correlated with $\Delta(p,\hat{\alpha}_i) = \pi(p,\hat{\alpha}_i) - p$ ($p < 0.005$, $k > 5$).

Method II: Fixed α^*

An alternative method is to use $\Delta(p,\alpha_i) = \pi(p,\alpha^*) - p$. Here, instead of using the best fit α for each subject, we use a single α value (I will choose the mean α, but this does not change the results qualitatively).

This method has the advantage of making very precise predictions on the how the coefficients b_1 and b_2 will change as the true α_i changes. For the b_1 term, this predicts that there will be a quadratic relationship between b_1 and α_i. Figure 49 shows this relationship for the individual α parameters observed in our sample (blue points), as well as all intermediate points (black line). In particular, note that the intercept is greater than 0 (as decision weight is strictly increasing in probability), and the quadratic relationship is exhibited by the inverse U-shape.

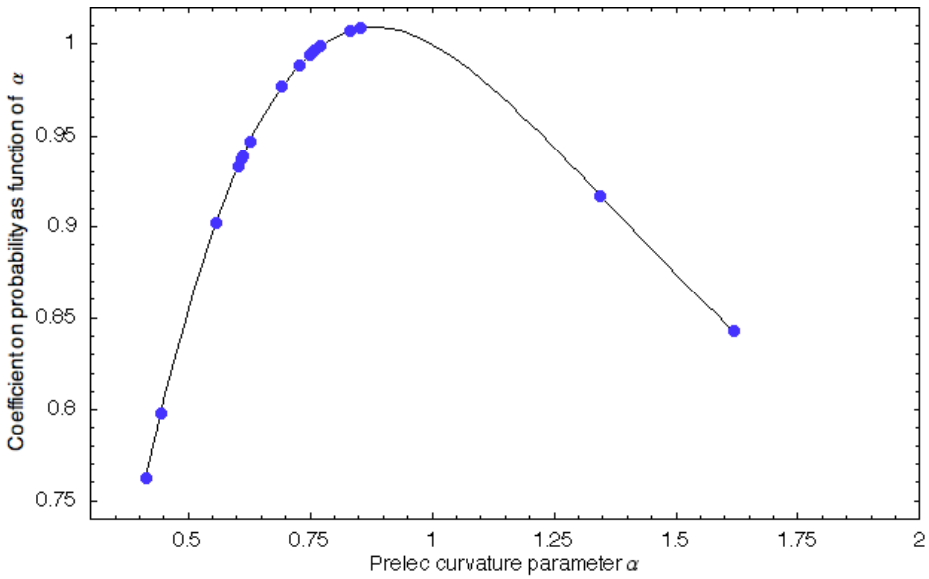


Figure 49: Relationship between regression coefficient b_1 *and* true α_i.

For the b_2 term, this predicts that there should be a negative relationship between b_2 and α_i. Figure 50 shows this relationship for the individual α parameters observed in our sample (blue points), as well as all intermediate points (black line). Furthermore, there is no appreciable nonlinearity in the values of α_i that we observed.

Another property of note is that the coefficient switches sign (from positive to negative) as it becomes greater than 1. This reflects the fact that people moves from underweighting of small probabilities to overweighting of small probabilities as α_i crosses unity (the pattern reverses for large probabilities, as expected). Overweighting of small probabilities is rare in the literature, and in our sample only two subjects are found to do this. Despite this, a combination of the fact that there exists a significantly negative relationship between b_2 and α_i, the regression model crosses 0 at approximately $\alpha_i = 1$, and significantly negative values for the two subjects with $\alpha_i > 1$, should be considered persuasive evidence.

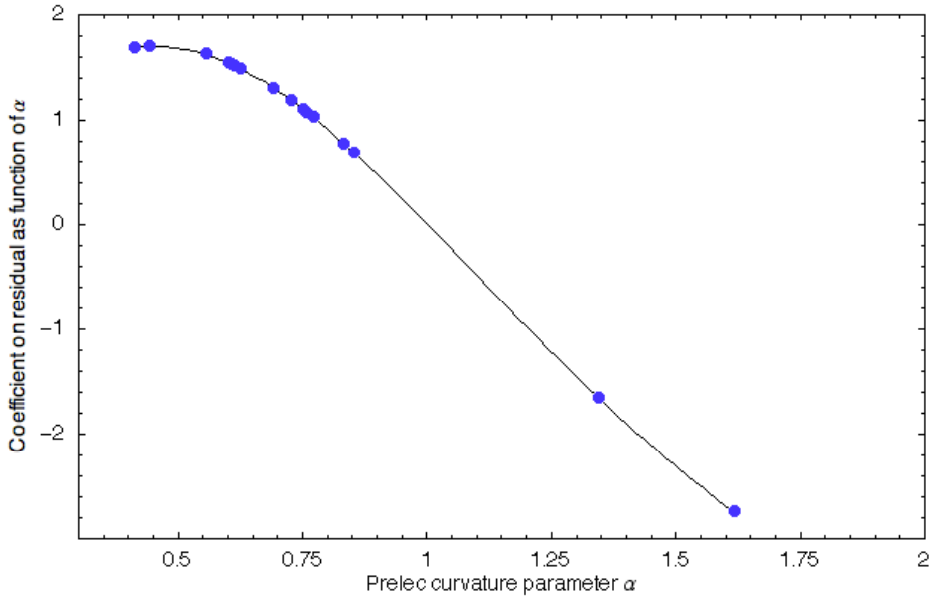


Figure 50: Relationship between regression term b_2 and true α_i.

We first look for regions that obey the predictions on the b_1 term. These are areas in which we regress the estimated b_1 coefficient for each subject on a quadratic $b_1 = \delta_1 + \delta_2\Delta(p,\alpha^*) + \delta_3\Delta(p,\alpha^*)^2$. A region is said to be active if it reflects all three of the predicted properties. That is, $\delta_1 > 0$, $\delta_2 < 0$, and $\delta_3 < 0$. This is a highly stringent criteria, and a very limited set of regions exhibit this combination of activity Table 25. In particular, Figure 51 shows a region near the caudate/insula that satisfies this conjunction analysis at $p < 0.05$ for each of the terms individually. Figure 58 shows the relationship between the estimated b_1's and α_i.

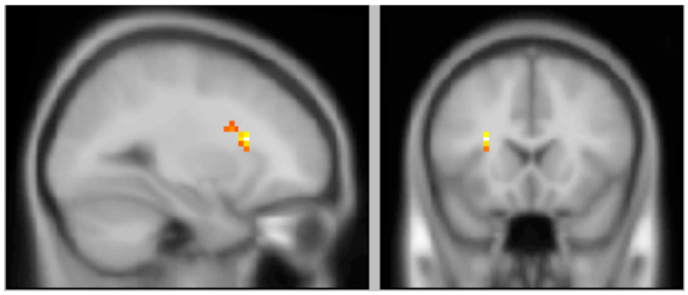


Figure 51: Activation of L caudate/insula (y = 18) in conjunction of constant, linear term, and quadratic term ($p < 0.05$ for each).

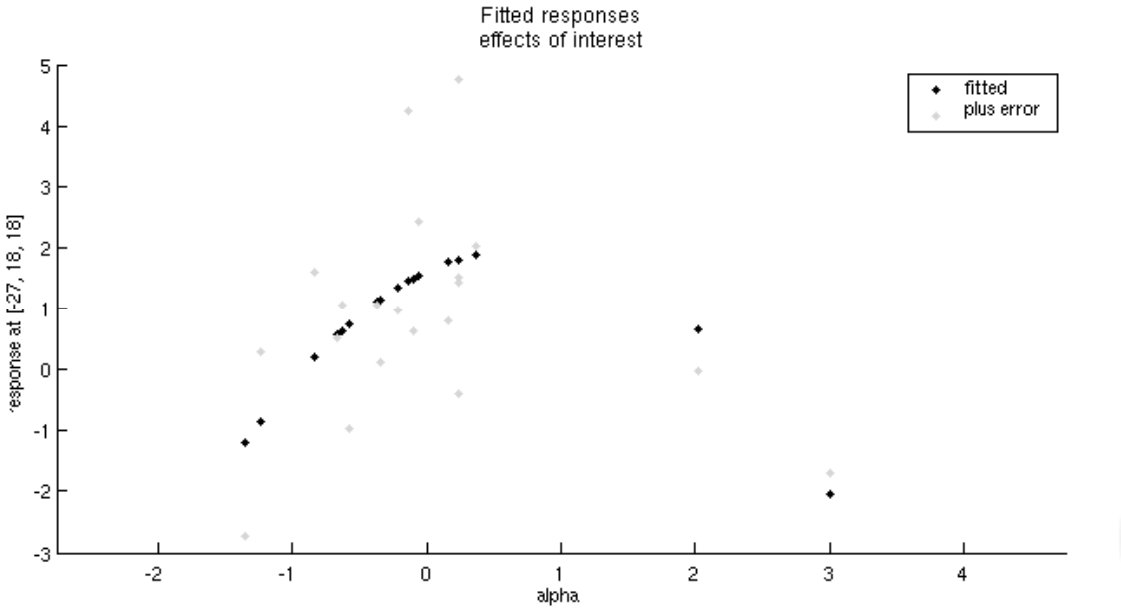


Figure 52: Fitted response of L caudate/insula (-27, 18, 18) as a function of estimated alpha.

k	p_{FWE}	p_{FDR}	T	p_{unc}	X	Y	Z	L/R	Region
14	0	0.152	30.91	0	-27	18	18	L	Caudate/Insula
6	0	0.152	12.49	0	63	-15	-18		
9	0	0.152	12.22	0	-9	-63	-15	R	Middle Temporal Gyrus

Table 25: Regions in conjunction analysis of areas that are significantly greater than 0, show negative linear and quadratic relationship alpha ($p < 0.05$ for each, $k > 5$).

Next we search for regions that obey the predictions on the b_2 term. These are areas in which we regress the estimated b_2 coefficient for each subject on $b_1 = \gamma_1 + \gamma_2 \Delta(p,\alpha^*) + \varepsilon$. A region is said to be active if it reflects both of the predicted properties. That is, $\gamma_1 > 0$, $\gamma_2 < 0$. This is a less stringent criterion than the previous, and a greater set of regions exhibit this combination of activity (Table 26). In particular, Figure 53 shows both the area activated in Figure 51, as well as a contralateral caudate region, that satisfies this conjunction analysis at $p < 0.04$ for each of the terms individually. Figure 52 shows the relationship between the estimated b_1's and α_i.

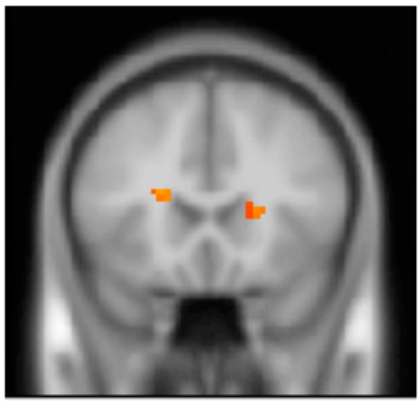


Figure 53: Conjunction of the L and R caudate (y = 18) in regression b_2 ($p < 0.04$).

The fitted responses shown in Figure 54 is notable in the fact that it fulfills all of the predictions mentioned in page 112 in reference to the regression model that produced Figure 50. Specifically, that there is a negative relationship between b_2 (this is true with or without the two subjects with $\alpha > 1$), that the regression line crosses 0 at approximately $\alpha = 1$, and the two subjects with $\alpha > 1$ do in fact have $b_2 < 1$ (although significance is difficult to verify).

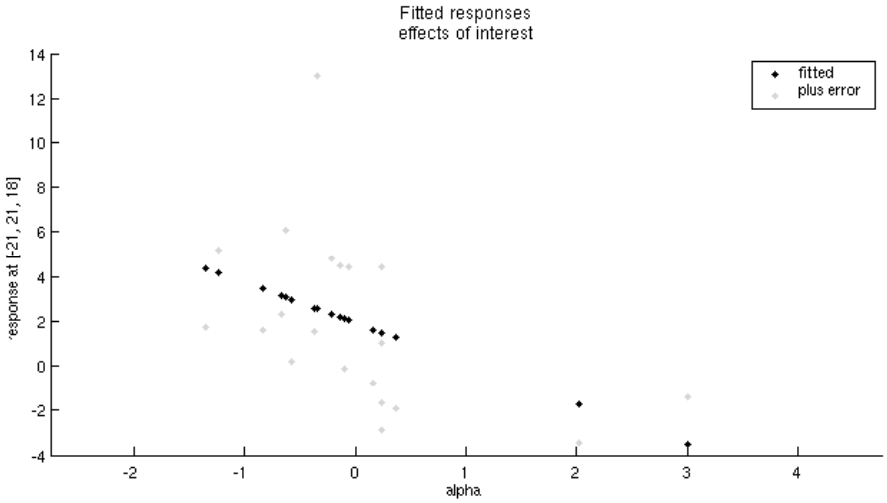


Figure 54: Fitted response of L caudate/insula (-21, 21, 18) as a function of estimated alpha.

k	p_{FWE}	p_{FDR}	T	p_{unc}	X	Y	Z	L/R	Region
65	0	0.308	11.85	0	33	-48	6	R	Parahippocampal gyrus/caudate tail

30	0.005	0.308	7.98	0	-36	-78	3	L	Middle Occipital Gyrus
10	0.012	0.308	7.46	0	-48	-18	-21	L	Middle Temporal Gyrus
12	0.223	0.308	5.57	0	24	12	54	R	Superior Frontal Gyrus
7	0.242	0.308	5.52	0	51	-51	-6	R	Middle Temporal Gyrus
19	0.246	0.308	5.5	0	-21	21	18	L	Insula/Caudate
7	0.289	0.308	5.38	0	3	3	3	R	Caudate head
12	0.31	0.308	5.33	0	-18	-27	15	L	Pulvinar/caudate tail
5	0.386	0.308	5.16	0	-30	-18	27	L	Postcentral gyrus
11	0.466	0.308	5	0	21	21	9	R	Caudate
9	0.534	0.308	4.88	0	-18	-90	0	L	Middle Occipital Gyrus
5	0.855	0.308	4.3	0	-9	-78	-27	L	Declive

Table 26: Regions in conjunction analysis of areas that are significantly greater than 0 and show decreasing linear relationship with alpha ($p < 0.04$, $k > 5$).

Five-Fold Pattern

Finally we search for regions that obey the predictions on the b_1 and b_2 terms. These are regions where $\delta_1 > 0$, $\delta_2 < 0$, $\delta_3 < 0$, $\gamma_1 > 0$, $\gamma_2 < 0$ each at $p < 0.1$. This is a very stringent criterion, and a small set of regions exhibit this combination of activity (Table 27). In particular, Figure 55 shows the caudate/insula region activated in Figure 51.

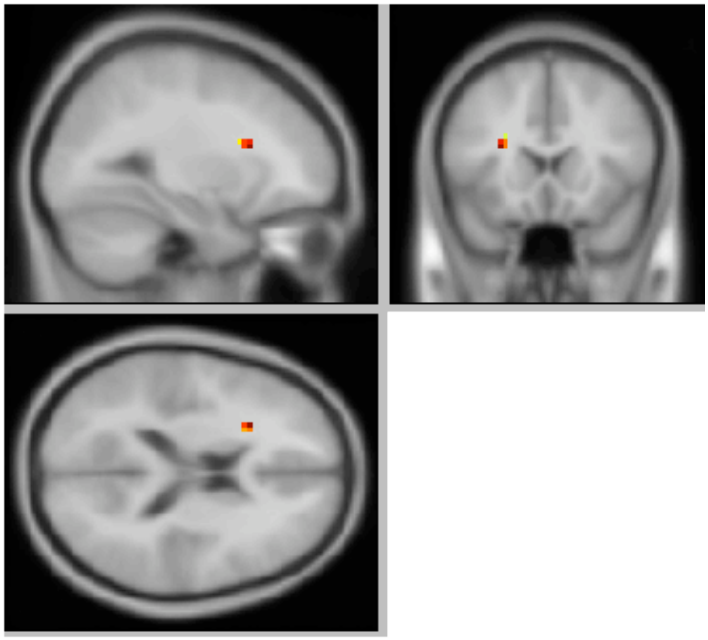


Figure 55: Activation with 5-fold pattern in caudate/insula (-27, 18, 18).

k	p_{FWE}	p_{FDR}	T	p_{unc}	X	Y	Z	L/R	Region
11	0	0.152	74.02	0	-27	18	18	L	Caudate/Insula
5	0	0.152	64.02	0	-21	-6	42	L	Middle Frontal Gyrus
13	0	0.152	36.76	0	-6	-63	-15	M	Culmen
5	0	0.152	13.23	0	-33	-81	-3	L	Middle Occipital Gyrus

Table 27: Regions in conjunction analysis of areas that show all five relationships ($p < 0.1$ for each, $k > 5$).

Encoding of Higher Moments

The encoding of the statistical moments other than the mean has only recently been studied in neuroscience. Even there, most studies have focused on the first two moments—mean and variance (Fiorillo et al. 2003; Preuschoff et al. 2006). Here, I present some evidence that the brain encodes for skewness and kurtosis as well (Figure 56).

First I note the correlations within the independent variables—that is, the moments. Because this study used binary gambles, such correlations are inevitable. This therefore limits our power in distinguishing between regions that possibly encode these moments. Table 15 shows the correlation between these measures as used in the experiment. The correlation between probability and gamble stake is zero by construction.

	prob	*stake*	*EV*	*Variance*	*Skew*	*Kurtosis*
Probability	1					
Gamble	0	1				
EV	0.64	0.61	1			
Variance	0.07	0.74	0.53	1		
Skewness	-0.43	0.06	-0.51	0.11	1	
Kurtosis	0.08	0.82	0.60	0.88	0.06	1

Table 28: Correlation between independent variables used in the experiment.

As is clear in the table, there is a substantial degree of correlation between some of the variables, notably between the variance and kurtosis, as well as (negative correlation) between the expected value and the skewness. It is therefore difficult to clearly distinguish between the expected value regions and the skewness regions, as well as variance regions between kurtosis regions.

Because of encoding of the first two moments, mean and variance, is relatively well-established, I will therefore construct the analysis to bias *against* finding the higher moments—that is, skewness and kurtosis. That is, we orthogonalized the independent variables in the following order via Gram-Schmidt in the following order: expected value, variance, skewness, and kurtosis. Therefore, skewness and kurtosis are effectively the residuals of the projection on expected value and variance (note that the correlation between skewness and kurtosis is close to 0). The brain regions found to be correlated with skewness and kurtosis, therefore, are areas that the skewness and kurtosis measures are explaining above and beyond those explained by variance and skewness.

Notably, these brain regions are correlated with the *raw* moments, and not the normalized moments. Interestingly, this is the Taylor expansion method that is used in finance to calculate the utility function.

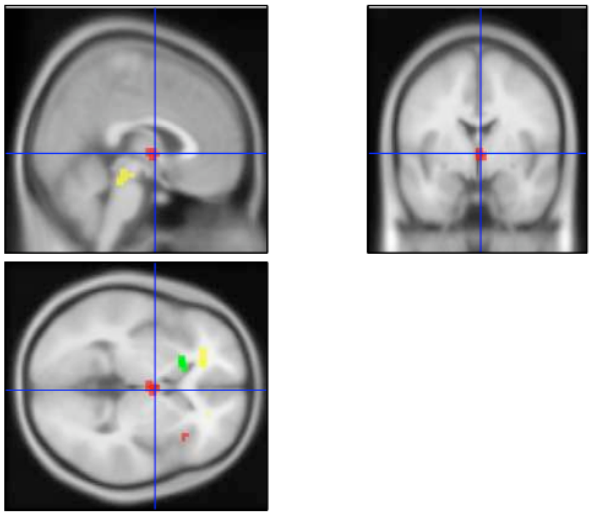


Figure 56: Regions significantly correlated with variance (caudate body and head, red), skewness (midbrain and caudate head, yellow), and kurtosis (caudate head, green).

Choice Regions

Thus far all data analyses on the neuroimaging data have been concentrated on the presentation of the first gamble. This is mainly due to the fact that the presentation of the first gamble allows us the cleanest look at how the brain encodes probabilities. In this

section we shall look at what happens in the brain during choice. We use the econometric model

$$y_i = \alpha_i + \beta_{i,1} EU(p_1, x_1; \alpha_i, \rho_i) + \beta_2 \Delta(\alpha_i, \rho_i) + \pi M_i + \varepsilon_i$$

where $EU(p_1, x_1; \alpha_i, \rho_i)$ is the expected utility of the first gamble given probability weighting parameter α_i and power utility parameter ρ_i for subject i, $\Delta(\alpha_i, \rho_i) = \left| EU(p_1, x_1; \alpha_i, \rho_i) - EU(p_2, x_2; \alpha_i, \rho_i) \right|$ is the absolute difference between the utility of the two alternatives, and π is the vector of nuisance parameters with associated matrix M.

The regions presented below in Table 29 are those that are significantly correlated with the (Prelec) expected utility of subjects' choices. That is, regions where the group $\beta_1 > 0$ in the above equation. Expected utilities are calculated using parameters derived from the stochastic choice model described on page 100 (both decision weights and utility (power) functions). Notable regions include caudate as well as the anterior cingulated (Figure 57).

k	p_{FWE}	p_{FDR}	*T*	p_{unc}	*X*	*Y*	*Z*	*L/R*	*Region*
973	0.11	0.017	5.81	0	27	-84	-12	R	Inferior Occipital Gyrus
	0.155	0.017	5.59	0	-9	-81	3	L	
	0.165	0.017	5.55	0	-9	-93	-3	L	
14	0.478	0.017	4.78	0	-39	-60	24	L	Brodman 39
16	0.56	0.017	4.64	0	-12	60	9	L	Superior Frontal Gyrus
35	0.652	0.019	4.49	0	-6	-3	-6	L	Caudate
	0.739	0.02	4.34	0	-9	9	0	L	
36	0.674	0.019	4.45	0	-54	-45	-12	L	Middle Temporal Gyrus
	0.945	0.027	3.86	0.001	-60	-48	-18	L	
30	0.711	0.019	4.39	0	9	12	3	R	Caudate
11	0.725	0.02	4.36	0	-3	30	-6	L	Anterior Cingulate
12	0.878	0.024	4.06	0	57	-36	-15	R	Middle Temporal Gyrus

Table 29: Regions significantly correlated with expected utility of choices ($p < 0.001$, $k > 10$).

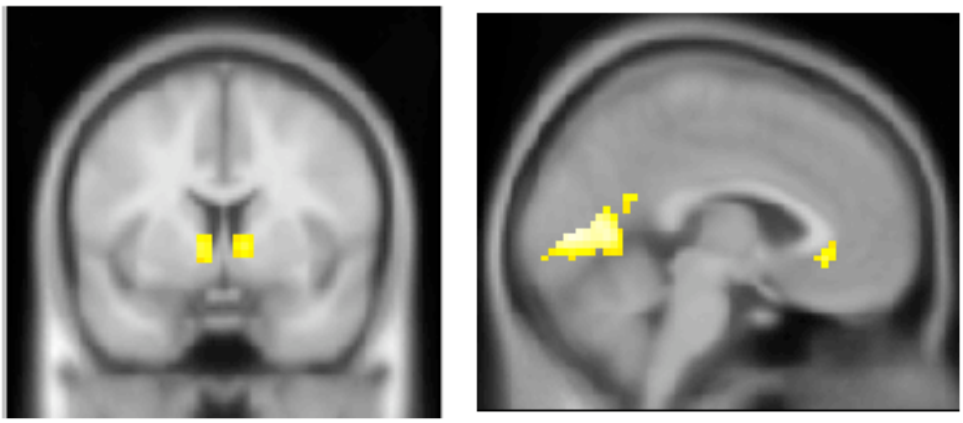


Figure 57: Regions including caudate (y = -3) and ACC (x = -3) are correlated with expected utility of choices ($p < 0.001$, $k > 10$).

Conflict Regions

A number of previous studies have found that areas such as the ACC are correlated with the "difficulty" of the choice (Kennerley et al. 2006). This is supported by the idea that the ACC is involved in action selection or executive conflict control (Etkin et al. 2006). We therefore look for regions that are negatively correlated with the absolute difference between the two choices. That is, regions where the hypothesis that group $\beta_2 = 0$ can be rejected. We base this on the hypothesis that the difficulty of the choices is greater when expected utility of the alternatives are close together.

The alternatives we used, however, were not optimized for fMRI analysis, but rather to be able to estimate behavioral parameters. Therefore, the expected values of the alternatives are highly correlated, and hence expected utility as well. Figure 59 shows the correlation between activation of the expected value of Gamble 1 and the absolute value of the difference between Gamble 1 and Gamble 2. As in previous analyses, we orthogonalize the absolute difference on the expected values, so that we bias our results *against* finding regions that encode conflict. Only 3 regions were found to be significantly negatively correlated with the difference in expected utility at a *very* liberal p-value of 0.01 and cluster size $k > 4$ (Table 30).

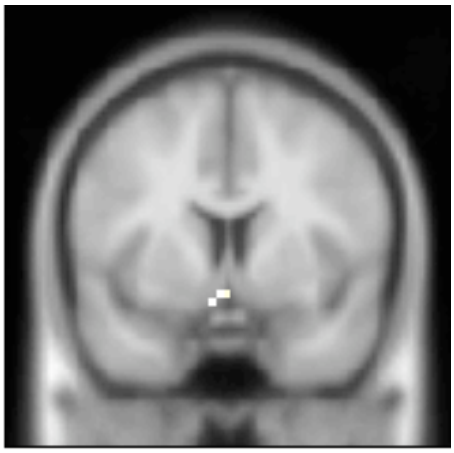


Figure 58: Ventral striatum/ACC (-3, 6, -12) is correlated with conflict at the very liberal p-value of 0.01 ($k > 4$ voxels).

As we noted above, because of the high degree of correlation between the absolute difference and the expected value of the alternatives, this negative result should be taken with some care. Indeed, the ACC activation we find in Figure 57 could well be due to the conflict measure.

k	p_{FWE}	p_{FDR}	T	p_{unc}	X	Y	Z	L/R	Region
21	0.997	0.96	3.08	0.003	63	-3	3	R	Superior Frontal Gyrus
5	0.998	0.96	3.07	0.003	-3	6	-12	L	Nucleus accumbens/ACC
5	1	0.96	2.66	0.008	-36	-84	6	L	Middle Occipital Gyrus

Table 30: Regions negatively correlated with absolute difference between the alternatives ($p < 0.01$, $k > 4$ voxels).

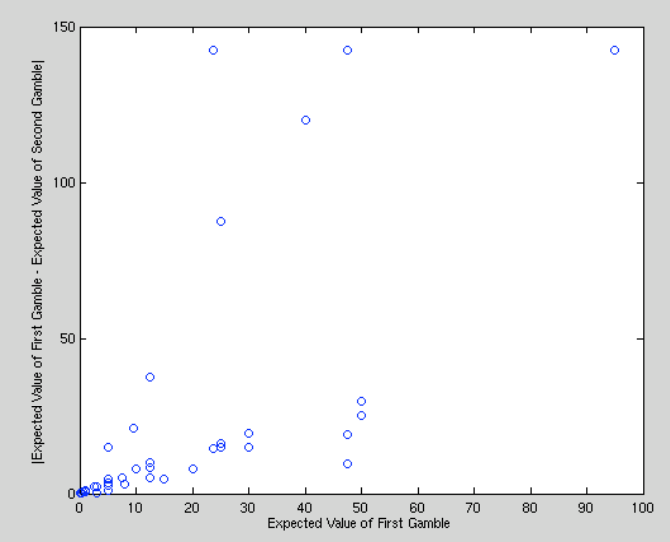


Figure 59: Correlation between expected value and absolute difference in expected value.

Fano Factor of Magnitude and Variance

As discussed in the statistical process of neuronal firing rates on page 3, the firing rate of neurons can be approximated by a Poisson distribution. The relation between the magnitude and variance of the firing should, therefore, be a linear one. The slope of this relationship is called the Fano Factor. It is clear by inspection of Figure 46 that there is a relationship between the magnitude of the beta and the standard error of the estimate. I estimated the relationship between the two in a log-log plot. Figure 60 shows this relationship for our data in our mixed effects model.

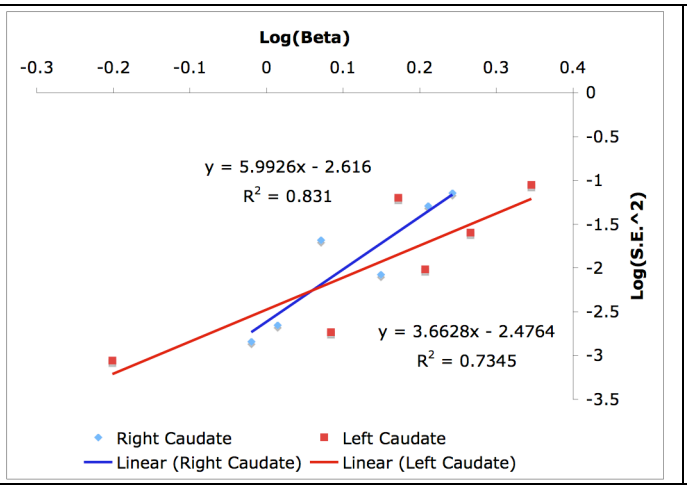


Figure 60: Scatter plot of log(beta) vs. log(variance). Because some of the betas were negative, a constant scaling factor were added to ensure positive betas.

Discussion

One brain region notably absent in our analysis is the LIP. As discussed in the introduction, the LIP is known to be involved in numerical processing, especially numerosity comparisons. The lack of finding may be due to the relatively high rate of false negatives in fMRI data. A more likely, and interesting, hypothesis, however, is that the LIP does encode the numerical quantities and comparisons in our task, but the simultaneous presentation of probabilities and outcomes meant that the different neural signals were confounded.

This chapter discusses a long-standing paradox in utility theory—Allais's Paradox. The primary goal of this chapter was to find the neural correlate of the empirical nonlinear weighting of probability. To this extent it is a partial success. Something like the probability weighting function found in economics was found, in the region that we had hypothesized. Yet it is clear, especially from the background discussion of the psychophysical data on proportion judgment, that *much* remains to be done.

This, in many ways, underscores the "interplay" between neuroscience and economics (as discussed by Dehaene with neuroscience and psychology in Chapter 1). This study started out as a fairly simple-minded search for the neural correlate of some behavioral regularity. Yet the data in neuroscience and psychophysics suggests that this phenomenon is the epiphenomenon of a much more intricate process of decision-making. Furthermore, these data point to very specific hypotheses and methods of testing them. All of this would not have been possible without taking the *science* behind decision sciences seriously.

Chapter 5

CONCLUSIONS

The goal of neuroeconomics is to ground economic theory in the details of how the brain works in decision-making, strategic thinking, and exchange. One method to achieve this is to observe processes and constructs that are typically considered unobservable, to decide between many theories of behavioral anomalies like risk aversion, altruistic punishment, and reciprocity, or to link processes (and data) from previously unconnected fields. This thesis aimed to contribute in both of these goals. Chapters 2 and 3 attempted to observe the neural substrates and the casual mechanisms, respectively, of ambiguity aversion. Chapter 4 attempted to observe the neural correlate of the probability weighting function, as well as to link a previously-unknown body of work by psychophysicists on the perception of proportion to probability weighting in decision-making under uncertainty.

As noted on page 7, a likely outcome of this interdisciplinary work between economics and neuroscience is that thinking about brain details will provide a new way of understanding concepts that have been traditionally left out of economic analysis, like emotion, willpower, habit, and the biological basis of demand. No one neuroeconomic study will be able to conclusively accomplish either of these things, but by combining experiments and types of data, we may gain insights into behavior that cannot be attained by introspection of behavioral observation.

The studies discussed in this paper illustrate all of these goals to varying degrees. First, with respect to decision-making under ambiguity, our data allows us to use neural evidence to discount models of ambiguity that treat is as merely a two-stage lottery, by showing that there may be two interacting systems determining the experienced utility of an option: one responding to the level of information in general, and one that uses input from this system to discount the reward of the possible results. Second, it suggests a substantially different

way of looking at risk-aversion. Considering a more general uncertainty aversion that dampens the utility of a gamble allows us to consider careful models of risk where the context of the gamble is important, something that is very difficult if we only consider a single universal utility function.

Separation of Perception and Choice

A theme throughout this thesis is that decision-making, as implemented by the brain, is likely to be separated into a number of stages. In particular, this predicts different neural mechanisms for the various stages of decision-making: perception, valuation, action selection, etc. The hypothesized separation of perception and choice can potentially present a unified explanation of the paradoxes studied in this thesis.

In addition to probability weighting and ambiguity aversion, research initiated by Heath and Tversky (1991) has shown that perceived competence can affect choice behavior—a phenomenon called "source preference." They found that perceived competence is inversely correlated with ambiguity aversion; that is, the weaker the perceived competence, the stronger the ambiguity aversion (Heath and Tversky 1991; Fox and Tversky 1995).

Some evidence for this was provided by Kilka and Weber (2001). They tested German subjects with gambles over German and Japanese bank stock returns, and then estimated separate (Lattimore) probability weighting functions for the stocks. They found that the Lattimore function had significantly greater curvature (γ), as well as lower elevation (δ) in the gambles over the Japanese (low competence) bank stock. Their findings are qualitative, as they only estimated the weighting function conditional on different utility function specifications (specifically, they used a power utility function with three different exponents at 0.76, 0.88, and 1).

Results from our study on both probability weighting and ambiguity aversion suggest that the effects of both are reflected in the dorsal striatum. To the extent that we can think of the dorsal striatum (and the caudate in particular) as the region that encodes reward value, this is as expected utility theory would suggest, however. It is likely the case that the

caudate represents the final product of a complex sequence of processing. Future steps should be to disentangle these sequences, which can potentially allow us to understand the mechanisms that generated the paradoxes that inspired the studies in this thesis.

BIBLIOGRAPHY

Adams, R. B., H. L. Gordon, A. A. Baird, N. Ambady and R. E. Kleck (2003). "Effects of gaze on amygdala sensitivity to anger and fear faces." Science **300**(5625): 1536-1536.

Alistair C Bruce, J. E. V. J. (2000). "Investigating the roots of the favourite-longshot bias: An analysis of decision making by supply- and demand-side agents in parallel betting markets." Journal of Behavioral Decision Making **13**(4): 413-430.

Allais, M. (1953). "Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'ecole americaine." Econometrica **21**(4): 503.

Amadi, A. (2004). Equity home bias: A disappearing phenomenon?, Available at SSRN: http://ssrn.com/abstract=540662.

Amaral, D., J. Price, A. Pitkanen and S. Carmichael (1992). Anatomical organization of the primate amygdaloid complex. The amydala: Neurobiological aspects of emotion, memory, and mental dysfunction. J. P. Aggleton. New York, Wiley.

Ashburner, J., K. Friston and W. Penny (2001). Human brain function, http://www.fil.ion.ucl.ac.uk/spm/doc/books/hbf2/.

Ashburner, J. and K. J. Friston (1999). "Nonlinear spatial normalization using basis functions." Human Brain Mapping **7**(4): 254-266.

Barberis, N. and M. Huang (2004). "Stocks as lotteries: The implications of probability weighting for security prices." AFA 2005 Philadelphia Meetings Paper. Available at SSRN: http://ssrn.com/abstract=649421 or DOI: 10.2139/ssrn.649421.

Bechara, A., D. Tranel, H. Damasio and A. R. Damasio (1996). "Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex." Cerebral Cortex **6**(2): 215-225.

Becker, S. W. and F. O. Brownson (1964). "What price ambiguity - or the role of ambiguity in decision-making." Journal of Political Economy **72**(1): 62-73.

Berthoz, S. (2002). "An fmri study of intentional and unintentional (embarrassing) violations of social norms." Brain **125**(8): 1696.

Bewley, T. (2002). "Knightian decision theory. Part i." Decisions in economics and finance **25**: 79 - 110.

Bleichrodt, H., J. van Rijn and M. Johannesson (1999). "Probability weighting and utility curvature in qaly-based decision making." Journal of Mathematical Psychology **43**(2): 238-260.

Boynton, G. M., J. B. Demb, G. H. Glover and D. J. Heeger (1999). "Neuronal basis of contrast discrimination." Vision Research **39**(2): 257-269.

Brusino, G. (1964). Note bibliographique sur le cours. Epstolario. V. Pareto. Rome, Accademia Nazionale dei Lincei**:** 1165-1172.

Caffarra, P. (1987). "Alexia without agraphia or hemianopia." European Neurology **27**(2): 65-71.

Camerer, C. (2003). Behavioral game theory : Experiments in strategic interaction. New York, N.Y., Princeton University Press.

Camerer, C. and R. Karjalainen (1994). Ambiguity-aversion and non-additive beliefs in non-cooperative games: Experimental evidence. Models and experiments on risk and rationality. B. Munier and M. Machina. Dordrecht, Kluwer Academic Publishers**:** 325-358.

Camerer, C., G. Loewenstein and D. Prelec (2005). "Neuroeconomics: How neuroscience can inform economics." Journal of Economic Literature **43**(1): 9-64.

Camerer, C. and M. Weber (1992). "Recent developments in modeling preferences - uncertainty and ambiguity." Journal of Risk and Uncertainty **5**(4): 325-370.

Chorvat, T. and K. Mccabe (2005). "Neuroeconomics and rationality." Chicago-Kent Law Review **80**: 101.

Churchland, P. S. (1986). Neurophilosophy: Toward a unified science of the mind-brain. Cambridge, Mass., MIT Press.

Churchland, P. S. and T. J. Sejnowski (1988). "Perspectives on cognitive neuroscience." Science **242**(4879): 741-745.

Critchley, H. D., C. J. Mathias and R. J. Dolan (2001). "Neural activity in the human brain relating to uncertainty and arousal during anticipation." Neuroimage **13**(6): S392-S392.

Dayan, P. and L. F. Abbott (2001). Theoretical neuroscience: Computational and mathematical modeling of neural systems, The MIT Press.

Dehaene, S. (2003). "The neural basis of the weber-fechner law: A logarithmic mental number." Trends in Cognitive Sciences **7**(4): 145-147.

Dehaene, S. and L. Cohen (1997). "Cerebral pathways for calculation: Double dissociation between rote verbal and quantitative knowledge of arithmetic." Cortex **33**(2): 219-250.

Dow, J. and S. R. D. Werlang (1992). "Uncertainty aversion, risk-aversion, and the optimal choice of portfolio." Econometrica **60**(1): 197-204.

Duff, P. (1999). "The scottish criminal jury: A very peculiar institution." Law and contemporary problems **62**(2): 173.

Ellsberg, D. (1961). "Risk, ambiguity, and the savage axioms." Quarterly Journal of Economics **75**(4): 643-669.

Epstein, L. G. and T. Wang (1994). "Intertemporal asset pricing under knightian uncertainty." Econometrica **62**(2): 283-322.

Erlick, D. E. (1964). "Absolute judgments of discrete quantities randomly distributed over time." Journal of Experimental Psychology **67**(5): 475-&.

Etkin, A., T. Egner, D. M. Peraza, E. R. Kandel and J. Hirsch (2006). "Resolving emotional conflict: A role for the rostral anterior cingulate cortex in modulating activity in the amygdala." Neuron **51**(6): 871-882.

Fiorillo, C. D., P. N. Tobler and W. Schultz (2003). "Discrete coding of reward probability and uncertainty by dopamine neurons." Science **299**(5614): 1898-1902.

Fox, C. R. and A. Tversky (1995). "Ambiguity aversion and comparative ignorance." Quarterly Journal of Economics **110**(3): 585-603.

Fox, J. (1997). Applied regression analysis, linear models, and related methods. Thousand Oaks, Calif., Sage Publications.

French, K. R. and J. M. Poterba (1991). "Investor diversification and international equity markets." American Economic Review **81**(2): 222-226.

Friedman, M. (1964). Essays in positive economics. Chicago ; London, University of Chicago Press.

Frisch, D. and J. Baron (1988). "Ambiguity and rationality." Journal of Behavioral Decision Making **1**: 149-157.

Friston, K. J., J. Ashburner, C. D. Frith, J. B. Poline, J. D. Heather and R. S. J. Frackowiak (1995a). "Spatial registration and normalization of images." Human Brain Mapping **3**(3): 165-189.

Friston, K. J., C. D. Frith, R. S. J. Frackowiak and R. Turner (1995b). "Characterizing dynamic brain responses with fmri - a multivariate approach." Neuroimage **2**(2): 166-172.

Friston, K. J., A. P. Holmes, K. Worsley, J. B. Poline, C. D. Frith and R. S. J. Frackowiak (1995c). "Statistical parametric maps in functional brain imaging: A general linear approach." Human Brain Mapping **2**: 189-210.

Friston, K. J., K. E. Stephan, T. E. Lund, A. Morcom and S. Kiebel (2005). "Mixed-effects and fmri studies." Neuroimage **24**(1): 244-252.

Fuortes, M. G. F. (1959). "Initiation of impulses in visual cells of limulus." Journal of Physiology-London **148**(1): 14-28.

Fuortes, M. G. F. and G. F. Poggio (1963). "Transient responses to sudden illumination in cells of the eye of limulus." J. Gen. Physiol. **46**(3): 435-452.

Gaffan, D., E. A. Murray and M. Fabrethorpe (1993). "Interaction of the amygdala with the frontal-lobe in reward memory." European Journal of Neuroscience **5**(7): 968-975.

Garcia, R., R. M. Vouimba, M. Baudry and R. F. Thompson (1999). "The amygdala modulates prefrontal cortex activity relative to conditioned fear." Nature **402**(6759): 294-296.

Garrett, T. A. and R. S. Sobel (1999). "Gamblers favor skewness, not risk: Further evidence from united states' lottery games." Economics Letters **63**(1): 85-90.

Ghirardato, P. and J. Katz (2005). "Indecision theory: Explaining selective abstension in multiple elections." Journal of Public Economic Theory **Forthcoming**.

Gilboa, I. and D. Schmeidler (1989). "Maxmin expected utility with non-unique prior." Journal of Mathematical Economics **18**(2): 141-153.

Glimcher, P. W. and A. Rustichini (2004). "Neuroeconomics: The consilience of brain and decision." Science **306**(5695): 447-452.

Glover, G. H. (1999). "Deconvolution of impulse response in event-related bold fmri." Neuroimage **9**(4): 416-429.

Golec, J. (1998). "Bettors love skewness, not risk, at the horse track." The journal of political economy **106**(1): 205.

Gonzalez, R. and G. Wu (1999). "On the shape of the probability weighting function." Cognitive Psychology **38**(1): 129-166.

Graham, J. R., C. R. Harvey and H. Huang (2006). Investor competence, trading frequency, and home bias, AFA 2006 Boston Meetings Paper Available at SSRN: http://ssrn.com/abstract=620801.

Gul, F. and W. Pesendorfer (2005). "The case for mindless economics." Princeton University Working Paper.

Gur, M., A. Beylin and D. M. Snodderly (1997). "Response variability of neurons in primary visual cortex (v1) of alert monkeys." Journal of Neuroscience **17**(8): 2914-2920.

Handwerker, D. A., J. M. Ollinger and M. D'Esposito (2004). "Variation of bold hemodynamic responses across subjects and brain regions and their effects on statistical analyses." Neuroimage **21**(4): 1639-1651.

Hardcastle, V. G. and C. M. Stewart (2002). "What do brain data really show?" Philosophy of Science **69**(3): S72-S82.

Harlow, J. M. (1869). Recovery from the passage of an iron bar through the head.

Heath, C. and A. Tversky (1991). "Preference and belief - ambiguity and competence in choice under uncertainty." Journal of Risk and Uncertainty **4**(1): 5-28.

Heeger, D. J., A. C. Huk, W. S. Geisler and D. G. Albrecht (2000). "Spikes versus bold: What does neuroimaging tell us about neuronal activity?" Nature Neuroscience **3**(7): 631-633.

Henson, R. (2001). Analysis of fmri timeseries: Linear time-invariant models, event-related fmri and optimal experimental design. Human brain function. J. Ashburner, K. Friston and W. Penny, http://www.fil.ion.ucl.ac.uk/spm/doc/books/hbf2/.

Hertwig, R., G. Barron, E. U. Weber and I. Erev (2004). "Decisions from experience and the effect of rare events in risky choice." Psychological Science **15**(8): 534-539.

Hollands, J. G. and B. P. Dyre (2000). "Bias in proportion judgments: The cyclical power model." Psychological review **107**(3): 500-524.

Hornak, J. P. (1997). Basics of mri: A hypertext book on magnetic resonance imaging.

Hubbard, E. M., M. Piazza, P. Pinel and S. Dehaene (2005). "Interactions between number and space in parietal cortex." Nature Reviews Neuroscience **6**(6): 435-448.

Hubel, D. H. and T. N. Wiesel (1959). "Receptive fields of single neurones in the cats striate cortex." Journal of Physiology-London **148**(3): 574-591.

Huettel, S. A., A. W. Song and G. McCarthy (2004). Functional magnetic resonance imaging. Sunderland, Mass., Sinauer Associates, Publishers.

Jevons, W. S. (1970). The theory of political economy. Harmondsworth,, Penguin.

Kahneman, D. and A. Tversky (1979). "Prospect theory - analysis of decision under risk." Econometrica **47**(2): 263-291.

Karmarkar, U. S. (1978). "Subjectively weighted utility: A descriptive extension of the expected utility model." Organizational behavior and human performance **21**(1): 69.

Kennerley, S. W., M. E. Walton, T. E. J. Behrens, M. J. Buckley and M. F. S. Rushworth (2006). "Optimal decision making and the anterior cingulate cortex." Nat Neurosci **9**(7): 940-947.

Kilka, M. (2001). "What determines the shape of the probability weighting function under uncertainty?" Management science **47**(12): 1712.

Knutson, B., C. Adams, G. Fong, J. Walker and D. Hommer (2001). "Parametric fmri confirms selective recruitment of nucleus accumbens during anticipation of monetary reward." Neuroimage **13**(6): S430-S430.

Kosfeld, M., M. Heinrichs, P. J. Zak, U. Fischbacher and E. Fehr (2005). "Oxytocin increases trust in humans." Nature **435**(7042): 673-676.

Kraus, A. (1976). "Skewness preference and the valuation of risk assets." The Journal of finance **31**(4): 1085.

Labar, K. S., J. E. Ledoux, D. D. Spencer and E. A. Phelps (1995). "Impaired fear conditioning following unilateral temporal lobectomy in humans." Journal of Neuroscience **15**(10): 6846-6855.

Lattimore, P. K., J. R. Baker and A. D. Witte (1992). "The influence of probability on risky choice - a parametric examination." Journal of Economic Behavior & Organization **17**(3): 377-400.

Lo, K. C. (1999). "Extensive form games with uncertainty averse players." Games and Economic Behavior **28**(2): 256-270.

Logothetis, N. K., J. Pauls, M. Augath, T. Trinath and A. Oeltermann (2001). "Neurophysiological investigation of the basis of the fmri signal." Nature **412**(6843): 150-157.

MacCrimmon, K. (1968). Descriptive and normative implications of the decision-theory postulates. Risk and uncertainty. K. Borch and J. Mossin. London, MacMillan**:** 3-32.

McClure, S. M., D. I. Laibson, G. Loewenstein and J. D. Cohen (2004). "Separate neural systems value immediate and delayed monetary rewards." Science **306**(5695): 503-507.

Mimura, M., R. Oeda and M. Kawamura (2006). "Impaired decision-making in parkinson's disease." Prarkinsonism Related Disorder **12**(3): 169-175.

Morse, A. and S. Shive (2006). Patriotism in your portfolio, Available at SSRN: http://ssrn.com/abstract=406200.

Mukerji, S. (1998). "Ambiguity aversion and incompleteness of contractual form." American Economic Review **88**(5): 1207-1231.

Mukherji, S. and J. M. Tallon (2004). An overview of economic applications of david schmeidler's models of decision

making under uncertainty. Uncertainty in economic theory: A collection of essays in honor of david schmeidler's 65th birthday. I. Gilboa, Routledge Publishers.

Nakajima, Y. (1987). "A model of empty duration perception." Perception **16**(4): 485-520.

Nelder, J. A. and R. Mead (1965). "A simplex-method for function minimization." Computer Journal **7**(4): 308-313.

Nieder, A. (2005). "Counting on neurons: The neurobiology of numerical competence." Nature Reviews Neuroscience **6**(3): 1-14.

Nieder, A. and E. K. Miller (2003). "Coding of cognitive magnitude: Compressed scaling of numerical information in the primate prefrontal cortex." Neuron **37**(1): 149-157.

O'Doherty, J. (2003). "Can't learn without you: Predictive value coding in orbitofrontal cortex requires the basolateral amygdala." Neuron **39**(5): 731-733.

O'Doherty, J., H. Critchley, R. Deichmann and R. J. Dolan (2003). "Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices." Journal of Neuroscience **23**(21): 7931-7939.

O'Doherty, J., P. Dayan, J. Schultz, R. Deichmann, K. Friston and R. J. Dolan (2004). "Dissociable roles of ventral and dorsal striatum in instrumental conditioning." Science **304**(5669): 452-454.

Ogawa, S., T. M. Lee, A. S. Nayak and P. Glynn (1990). "Oxygenation-sensitive contrast in magnetic-resonance image of rodent brain at high magnetic-fields." Magnetic Resonance in Medicine **14**(1): 68-78.

Ongur, D. and J. L. Price (2000). "The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans." Cerebral Cortex **10**(3): 206-219.

Patterson, J. C., L. G. Ungerleider and P. A. Bandettini (2002). "Task-independent functional brain activity correlation with skin conductance changes: An fmri study." Neuroimage **17**(4): 1797-1806.

Paulus, M. P., C. Rogalsky, A. Simmons, J. S. Feinstein and M. B. Stein (2003). "Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism." Neuroimage **19**(4): 1439-1448.

Pessiglione, M., V. Czernecki, B. Pillon, B. Dubois, M. Schupbach, Y. Agid and L. Tremblay (2005). "An effect of dopamine depletion on decision-making: The temporal coupling of deliberation and execution." J. Cogn. Neurosci. **17**(12): 1886-1896.

Phelps, E. A., M. R. Delgado, K. I. Nearing and J. E. LeDoux (2004). "Extinction learning in humans: Role of the amygdala and vmpfc." Neuron **43**(6): 897-905.

Phelps, E. A., K. J. O'Connor, W. A. Cunningham, E. S. Funayama, J. C. Gatenby, J. C. Gore and M. R. Banaji (2000). "Performance on indirect measures of race evaluation predicts amygdala activation." Journal of Cognitive Neuroscience **12**(5): 729-738.

Phelps, E. A., K. J. O'Connor, J. C. Gatenby, J. C. Gore, C. Grillon and M. Davis (2001). "Activation of the left amygdala to a cognitive representation of fear." Nature Neuroscience **4**(4): 437-441.

Pinheiro, J. C. and D. M. Bates (2000). Mixed-effects models in s and s-plus. New York, Springer.

Platt, M. L. and P. W. Glimcher (1999). "Neural correlates of decision variables in parietal cortex." Nature **400**(6741): 233-238.

Prelec, D. (1998). "The probability weighting function." Econometrica **66**(3): 497-527.

Preuschoff, K., P. Bossaerts and S. Quartz (2006). "Encoding expected reward and risk in economic decision making." Neuron **Forthcoming**.

Quirk, G. J., J. L. Armony and J. E. LeDoux (1997). "Fear conditioning enhances different temporal components of tone-evoked spike trains in auditory cortex and lateral amygdala." Neuron **19**(3): 613-624.

Quirk, G. J., J. C. Repa and J. E. Ledoux (1995). "Fear conditioning enhances short-latency auditory responses of lateral amygdala neurons - parallel recordings in the freely behaving rat." Neuron **15**(5): 1029-1039.

Rees, G., K. Friston and C. Koch (2000). "A direct quantitative relationship between the functional properties of human and macaque v5." Nature Neuroscience **3**(7): 716-723.

Ress, D., B. T. Backus and D. J. Heeger (2000). "Activity in primary visual cortex predicts performance in a visual detection task." Nature Neuroscience **3**(9): 940-945.

Rieke, F. (1997). Spikes : Exploring the neural code. Cambridge, Mass., MIT Press.

Rogers, R. D., B. J. Everitt, A. Baldacchino, A. J. Blackshaw, R. Swainson, K. Wynne, N. B. Baker, J. Hunter, T. Carthy, E. Booker, M. London, J. F. W. Deakin, B. J. Sahakian and T. W. Robbins (1999). "Dissociable deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: Evidence for monoaminergic mechanisms." Neuropsychopharmacology **20**(4): 322-339.

Roy, C. S. and C. S. Sherrington (1890). "On the regulation of the blood supply of the brain." Journal of Physiology **11**: 85-108.

Rustichini, A., J. Dickhaut, P. Ghirardato, K. Smith and J. V. Pardo (2005). "A brain imaging study of the choice procedure." Games and Economic Behavior **52**(2): 257-282.

Sargent, T. and L. Hansen Robust control and economic model uncertainty, New York University.

Savage, L. J. (1954). The foundations of statistics. New York,, Wiley.

Schmeidler, D. (1989). "Subjective-probability and expected utility without additivity." Econometrica **57**(3): 571-587.

Schoenbaum, G., A. A. Chiba and M. Gallagher (2000). "Changes in functional connectivity in orbitofrontal cortex and basolateral amygdala during learning and reversal training." Journal of Neuroscience **20**(13): 5179-5189.

Schultz, W. (2000). "Multiple reward signals in the brain." Nature Reviews Neuroscience **1**(3): 199-207.

Schultz, W., P. Dayan and P. R. Montague (1997). "A neural substrate of prediction and reward." Science **275**(5306): 1593-1599.

Schultz, W., L. Tremblay and J. R. Hollerman (2000). "Reward processing in primate orbitofrontal cortex and basal ganglia." Cerebral Cortex **10**(3): 272-283.

Scoville, W. B. and B. Milner (1957). "Loss of recent memory after bilateral hippocampal lesions." Journal of Neurology Neurosurgery and Psychiatry **20**(1): 11-21.

Segal, U. (1987). "The ellsberg paradox and risk-aversion - an anticipated utility approach." International Economic Review **28**(1): 175-202.

Selden, N. R. W., B. J. Everitt, L. E. Jarrard and T. W. Robbins (1991). "Complementary roles for the amygdala and hippocampus in aversive-conditioning to explicit and contextual cues." Neuroscience **42**(2): 335-350.

Shulman, R. (1996). "Interview with robert g. Shulman." Journal of cognitive neuroscience **8**: 474-480.

Siegel, J. (1998). Stocks in the long run. New York, Irwin.

Spence, I. I. (1990). "Visual psychophysics of simple graphical elements." Journal of experimental psychology. human perception and performance **16**(4): 683-692.

Strong, N. and X. Z. Xu (2003). "Understanding the equity home bias: Evidence from survey data." Review of Economics and Statistics **85**(2): 307-312.

Sugrue, L. P., G. S. Corrado and W. T. Newsome (2005). "Choosing the greater of two goods: Neural currencies for valuation and decision making." Nature Reviews Neuroscience **6**(5): 363-375.

Tobler, P. N., C. D. Fiorillo and W. Schultz (2005). "Adaptive coding of reward value by dopamine neurons." Science **307**(5715): 1642-1645.

Tversky, A. and D. Kahneman (1992). "Advances in prospect-theory - cumulative representation of uncertainty." Journal of Risk and Uncertainty **5**(4): 297-323.

Varey, C. A., B. A. Mellers and M. H. Birnbaum (1990). "Judgments of proportions." Journal of Experimental Psychology-Human Perception and Performance **16**(3): 613-625.

Whalen, P. J. (1998). "Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala." Current Directions in Psychological Science **7**(6): 177-188.

Whalen, P. J., S. L. Rauch, N. L. Etcoff, S. C. McInerney, M. B. Lee and M. A. Jenike (1998). "Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge." Journal of Neuroscience **18**(1): 411-418.

Willingham, D. T. and E. W. Dunn (2003). "What neuroimaging and brain localization can do, cannot do, and should not do for social psychology." Journal of Personality and Social Psychology **85**(4): 662-671.

Wu, G. and R. Gonzalez (1996). "Curvature of the probability weighting function." Management Science **42**(12): 1676-1690.

Zak, P. J. (2004). "Neuroeconomics." Philosophical Transactions of the Royal Society of London Series B-Biological Sciences **359**(1451): 1737-1748.

Zangaladze, A., C. M. Epstein, S. T. Grafton and K. Sathian (1999). "Involvement of visual cortex in tactile discrimination of orientation." Nature **401**(6753): 587-590.

ENDNOTES

[1] However, Colander (2005)(23; 20) notes that Ramsey, Edgeworth and Fisher all speculated about measuring utility directly before the neoclassical revolution declared utility to be inherently unobservable and only revealed by observable choices. Ramsey and Edgeworth speculated about a "psychogalvanometer" and a "hedonimeter", respectively, which sound remarkably like modern tools. Would these economists be neuroeconomists if they were reincarnated today?

[2] The most charitable way to interpret Friedman's "F-twist" is that theories of market equilibrium based on utility-maximization of consumers are equivalent, at the market level, to unspecified other theories which allow violations of utility-maximization but include some institutional repairs or corrections for those violations. That is, the theory states $U \rightarrow M$, but even if U is false, there is an alternative theory [(*not - U)andR*] $\rightarrow M$, in which a "repair condition" R suffices in place of assumption U, to yield the same prediction M. If so, the focus of attention should be on specifying the repair condition R.

[3] "Not proven" is a vestige of an earlier time when juries just decided on whether facts were proven or not proven, and judges decided guilt based on those factual judgments.

[4] Returning to the Ellsberg example, a prior is simply the vector (p, 1 - p) where p is the probability that the state is red. As you can see the set of overall possible priors is the line segment in the real plane going between the points (0,1) - definitely black, to (1,0) -definitely red. A decision makers set of priors in this example is taken to be an unbroken segment of this line.

[5] In the Ellsberg example larger intervals for possible *p*

[6] The time series B_t for each voxel went through a high-pass filter and an AR(1) correction.

[7] To calculate the expected value of the gamble in the Informed Opponents treatment, we solve first the optimal strategy of the opponent. Denote (N,R,n,r) as respectively the number of cards total in the deck, the number of red cards total in the deck, the number of cards in the sample the opponent observes, and the number of red cards in the sample the opponent observes.

The opponent knows (N,n,r). His optimal strategy is therefore simply to choose the more likely color in the deck given the realizations of his sample. That is, the opponent chooses according to $\max\left(\frac{E(R \mid r,n,N)}{N}, 1-\frac{E(R \mid r,n,N)}{N}\right)$. The only unknown in the expression is the numerator. To solve, apply Bayes' rule, such that $P(R \mid r,n,N) = \frac{P(R \mid N)P(r \mid R,n,N)}{\sum_{i=0}^{N} P(r \mid i,n,N)}$, where $P(R \mid N)$ is the prior

distribution on *R*, $P(r \mid R,n,N)$ is the hypergeometric distribution (as this is an example of sampling without replacement), and the denominator is the probability of observing *r* over the support of *R*. With $P(R \mid r,n,N)$ in hand, the opponent can calculate the expected number of red cards in the deck, which is $E(R \mid r,n,N) = \sum_{R=0}^{N} P(R \mid r,n,N) \times R$.

Because this is a constant-sum game, the subject's probability of winning is the expected proportion of color that the subject is betting on in the deck $P(win \mid r,n,N) = \min\left(\frac{E(R \mid r,n,N)}{N}, 1 - \frac{E(R \mid r,n,N)}{N}\right)$. That is, the complement of the opponent's. As the subject does not observe *r*, we need to take the expecation over *r*. That is, $P(win \mid n,N) = \sum_{r=0}^{n} P(r \mid n,N) \times P(win \mid r,n,N)$.

Given the independence of the choices of the subject and the opponent, their choices will coincide with $p = 0.5$ in expectation. In this case, according to the rules of the game, the bet does not take place and both earn the certain payoff. The payoff function for the subject is therefore $0.5 \times P(win \mid n,N) \times x + 0.5 \times c$, where *x* is the amount of the gamble, and *c* is the certain payoff.

Finally, note that the choice of priors *P(R|N)* is left unspecified in the above. For example, if the subject believes that the deck is composed of either all red cards or all blue cards (*P(R=N|N)=0.5* and *P(R=0|N)=0.5*), a sample from the deck would determine completely the composition of the deck. We present in Table S3 expected value calculations given a uniform prior—$P(R \mid N) = 1/N + 1, \forall R$, and a binomial prior with $p = ½$—$P(R \mid N) = \binom{N}{R}\left(\frac{1}{2}\right)^{R}\left(\frac{1}{2}\right)^{N-R}$.

[8] (see the RC-AC image at http://www.econ.umn.edu/ arust/neuroecon.html).