CHAPTER

23

c00023

The Neurobiological Foundations of Valuation in Human Decision-making under Uncertainty

Peter Bossaerts, Kerstin Preuschoff and Ming Hsu

OUTLINE

s0010

INTRODUCTION

p0180 The goal of this chapter is to review recent neurobiological evidence to improve our understanding of human valuation under uncertainty. Although ultimately interested in human behavior, we will borrow from studies of animals with related brain structures, namely, non-human primates. Specifically, we wish to explore how valuation is accomplished. As we shall see, the evidence rejects a pure "retrieval from memory" model; instead, values are *computed*. This raises the issue: what computational model(s) are being used? Since actual choice can be summarized in terms of a single-dimensional utility index as in expected utility or prospect theory, we want to know how such an index is computed, and to understand the effect of perceptual biases on this computation, as well as the role of emotions. How does the computational model generate the risk aversion that we may see in choices? Or, in ambiguous situations, how is ambiguity aversion revealed in choices (Hsu *et al.*, 2005; Huettel, *et al.* 2006; Bali *et al.*, 2008) and what model underlies it – for example, alpha-maxmin preferences (Ghirardato *et al.*, 2004), anticipated regret (Segal, 1987), or some other?

p0190 A natural question for economists is, if choice can be represented "as if" some utility index is maximized, why should we bother studying the computational aspects behind choice? Our quest for the computational model underlying valuation has two purposes. First, we hope to be able to improve choice prediction. We consider the utility index model as a reduced-form model of choice, which, in analogy with reduced-form models of the macro-economy, is sensitive to circumstances – i.e., its parameters need to be revised every time one moves from one framework to another. This is the so-called *Lucas critique* (Lucas, 1976). This sensitivity, or lack of robustness, has been frequently noted in the literature as "framing effects." Our computational model, in contrast, is meant to be structural, and hence has a chance of outperforming the reduced-form utility index model. Of course, the proof is in the pudding …

p0200 A computational model should not only enable us to accommodate framing biases; it should also be capable of capturing part, if not all, of the randomness one usually adds to a utility-based choice model to accommodate the erratic nature of observed choice. The traditional approach in applied economics has been to consider the utility index itself to be random, and to attribute certain distributional properties to this randomness that ultimately leads to a choice-model based on logistic regression (McFadden, 1974). The goal of exploring computational models of choice is to provide a different (computational) basis for the erratic nature of observed choice.

p0210 Artificial intelligence is the second reason why we are interested in the actual neurobiological computational model behind observed choices. Modern decision theory has been quite successful in prescribing learning and choice in well-defined situations. Once the complexity of the problem increases, however, decision theory has little concrete to propose. One case in point is the Iowa Gambling Task (IGT) (Bechara and Damasio, 2005), which is a complex four-armed bandit problem. To date, no-one has been able to spell out the optimal strategy in this task. Humans, however, can quickly learn the ordering (in terms of reward and risk) of the four bandits (Bechara *et al*., 1997).

p0220 However, the learning model behind human choice in difficult situations such as the IGT is almost surely not Bayesian, despite the popularity of this approach. Open-ended situations, almost by definition, involve high-dimensional parameter spaces, and it is well known that Bayesian learning generically fails to converge when the parameter space becomes large, while simple learning algorithms based on classical statistics often do a very good job in capturing the essence of the stochasticity at hand (Diaconis and Freedman, 1986). This in turn leaves us the question, how do humans learn in such situations? What computational model are they using? If we knew this model, perhaps we could enrich decision theory to make it work in more open-ended or even ill-defined decision situations.

One clue in the search is provided by the plethora of brain lesion data, much of it using the IGT (despite problems with interpretability – see Dunn *et al*., 2006), supporting the idea that specific brain regions underpin these computations. Of particular interest are the ventromedial prefrontal cortex and the amygdala, as patients with lesions to these brain regions are impaired in the IGT while at the same time showing normal or even superior performance in working memory, IQ, and decisions and games that are less open-ended (Bechara *et al*., 1997, 1999). p0230

In this chapter, we focus on valuation in the context of choice under uncertainty. The computational model that emerges from a study of the neurobiological data is one where expected rewards are traded off against risk (or, in the case of risk-seeking behavior, risk is added to expected reward). This model has been popular in finance, where risk is usually represented in terms of reward variance (although more advanced models take into account higher moments) – hence its name: the *mean-variance model*. p0240

The link between the utility models representing actual choice under uncertainty, such as expected utility and prospect theory, on the one hand, and the mean-risk model, on the other, is often clarified by means of Taylor series expansions (Bali *et al*., 2008; see also Box 23.1). One of the goals of this chapter is to demonstrate that the logic of a computational model based on a tradeoff between expected reward and risk can be extended to choice under ambiguity as well. p0250

Throughout this chapter, we take the position that valuation stands on its own and makes sense even in the absence of choice. It makes sense because computing values take energy, and hence an organism that has learned to correctly compute the values of options it is forced to take on (imperative trials) will also be better able to determine the adequate choices when there is freedom (free-choice trials). As we shall see, brain activation when a subject is faced with gambles with uncertain outcomes is comparable across imperative and free-choice trials. p0260

When valuation is a pursuit distinct from choice, values revealed through choice may be different from the valuations that emerge from the (choice-independent) computations. So are economists interesting in knowing what the latter are? After all, economists are only interested in choice. We shall spend some time elaborating on this issue. Suffice it to say here that the distinction is important for welfare assessment. p0270

The remainder of this chapter is organized as follows. The following section elaborates on the distinction between values the brain computes and the values revealed through choice. We then review the neurobiological foundations of valuation under pure p0280 b0010 BOX 23.1

MEAN-VARIANCE UTILITY

p0010 The expected utility approach, in which the decision-maker maximizes a weighted sum of the utilities, dominates in economics. In finance, however, mean-variance (or risk-return) models provide the central approach to decision-making. In it, the decision-maker trades off between mean (return) and variance (risk). The two approaches are related. There are conditions under which the two models are identical, such as when returns are approximately normal, or if the utility function is quadratic (Markowitz, 1952). More generally, the two can be linked by assuming that the mean-variance approach is an approximation of the utility function through a Taylor series approximation.

p0020 Specifically, given a utility function $U(R)$, where R is the rate of return in the current period, we can implement a second-order Taylor series approximation around the mean $\mu = E(\mathrm{R})$, such that

p0030

$$U(R) \approx U(\mu) + U'(\mu)(R - \mu) + U''(R - \mu)^2.$$

p0040 The expected utility is therefore

p0050

$$E(U(R)) \approx U(\mu) + U''(\mu)\sigma^2,$$

p0060 where the second term $\sigma^2 = Var(R)$ is the variance of the return.

risk. The final section extends the logic of expected reward and risk to situations with ambiguity. We will not explore the neurobiological foundations of the translation of valuation into choice, as these are covered in later chapters in this volume.

s0020 DISTINGUISHING VALUATION AND CHOICE

p0290 In many human endeavors, valuation is performed even in the absence of any immediate necessity to choose. Again, finance is a case in point, in part because financial valuation is often complex and time-consuming, while good choice opportunities are rare and short-lived. The cost of computing values provides a normative rationale for why valuation may be done in the absence of free choice. Figure 23.1 provides strong evidence that valuation is performed even in imperative trials, and in a way that is relevant for free-choice trials. It shows how well a valuation model fit to neuronal activation during imperative trials can predict choice in free-choice trials. In fact, the neurobiological choice model predicts choices better than a utility-index based model estimated from the choices themselves! This demonstrates not only that valuation is done during imperative trials, but that the resulting values are relevant for choice in free-choice trials as well.

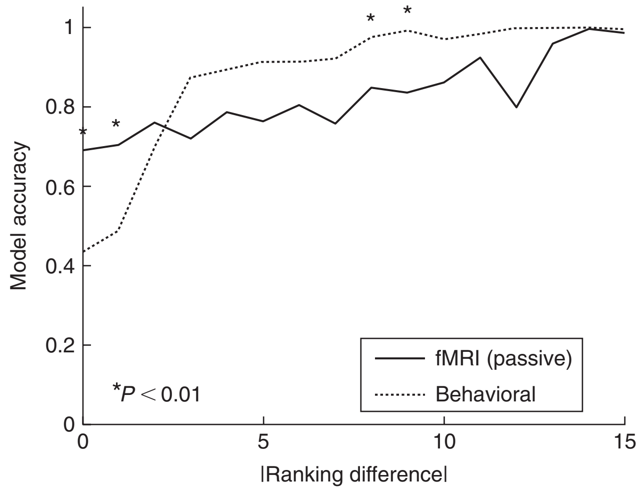


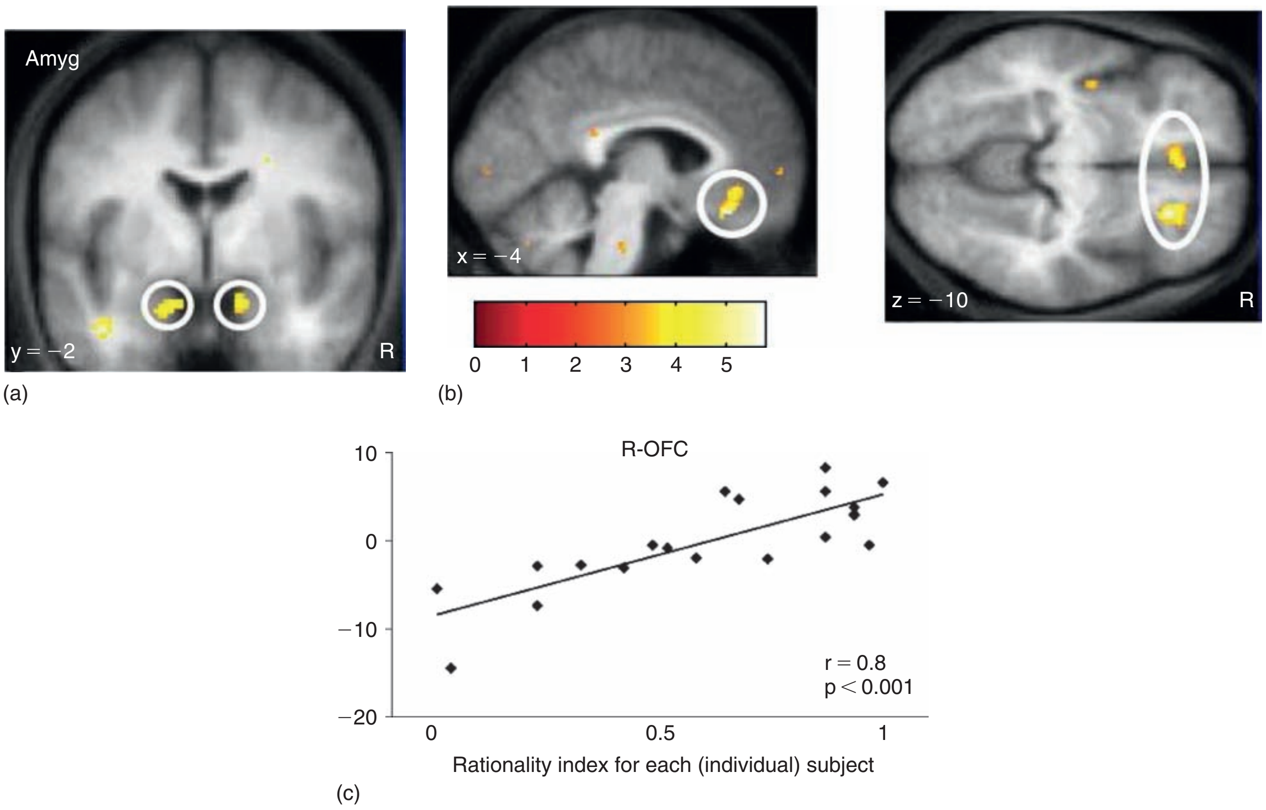
FIGURE 23.1 Out-of-sample prediction accuracy (percentage correct) of two models of binary choice, as a function of the ranking difference of the two options revealed through choice. The fMRI (passive) model uses as inputs activation in brain regions identified to be involved in valuation during purely imperative trials. The behavioral model is based on rankings revealed in choice in free-choice trials. The parameters of both models are fit on free-choice trials. Prediction accuracy is out-of-sample, i.e., it is measured in free-choice trials not used for fitting the models. Neural activation identified in imperative trials (the fMRI model) predicts choices better than actual choice in free-choice trials when the two available options are close in ranking. Adapted from Berns *et al.* (2007). f0010

p0300 Although brain activation during imperative trials reflects valuations that are compatible with the values revealed in free-choice trials, and hence that brain activation in imperative trials can be used to predict choice in free-choice trials (Figure 23.1), the fit is not 100%. These neurobiological data suggest that there are (at least) two value signals: one revealed through activation in brain regions not directly involved in the physical implementation of choice, and a second one revealed through activation of the neurons controlling the physical act of choice. Given economists' interest in choice, to what extent might they be interesting in knowing the former value signal?

p0310 The most important argument concerns welfare. What is it that the subject really wanted? Indeed, the mere presence of two (or more) valuation signals suggests that there may be a difference between what subjects wanted and what they did. We may never be able to determine what exactly subjects wanted (i.e., to extract their true preferences), but a better understanding of how the various value signals are put together may help us to better appreciate the difference between revealed and true preferences.

p0320 It is fundamental to question coincidence of revealed and true preferences (psychologists would distinguish between decision and experienced utility – see Kahneman *et al*., 1997). Is a utility-index model such as prospect theory a preference theory? Prospect theory is known to fit choice well, but do the values of its utility index reflect real preferences? The insistence on deriving choice-theoretic axiomatic foundations for prospect theory (Kahneman and Tversky, 1992) indicates that economists consider prospect theory to be a preference theory. However, the risk-seeking attitude towards losses that is implied by prospect theory, for instance, may not reflect context-adapted goal-oriented behavior, but a "default action" that is appropriate to maximize experienced utility only on average. The default action is robust to lapses of attention, and hence is optimal for an organism that has to spend effort to interpret the true meaning of stimuli around itself. It may be considered a "habit," but the neuroscience literature reserves the term *habit* for actions that are stimulus-sensitive but goal-insensitive (Yin *et al*., 2004). In contrast, default actions are stimulus-insensitive and goal-oriented. They are not instincts, either, because they can be learned.

p0330 Recent neurobiological evidence provides support for the notion of default actions. Choice in loss situations appears to be associated with activation in amygdala, irrespective of a subject's tendency to take on risk. Orbitofrontal cortex (OFC) activation, however, decreases with subjects' tendency to become risk-seeking with losses (Figure 23.2). These data suggest a very different valuation model underlying choice, based on default actions that need to be overcome in situations where these actions are inappropriate – yet not everyone manages to over-ride the default. In this interpretation of the data, amygdala activation engages the default action, while OFC is engaged in over-riding this choice when inappropriate. Perhaps no one wants to be risk-seeking *for* monetary losses,



f0020 FIGURE 23.2 (a) Amygdala activation correlates with contrast between lotteries cast in terms of gains and in terms of losses. (b, c) Contrast in prefrontal cortex activation between lotteries cast in terms of gains and in terms of losses increases with subjects' rationality index. This index is measured as the difference between the proportion of trials in which subjects chose the gamble in the loss frame, as compared to the gain frame, and then linearly transformed such that 1 is the most rational. Adapted from De Martino *et al*. (2006).

but many subjects cannot overcome their tendency to become risk-seeking under adverse conditions.

p0340 More generally, it is not so much the value signals that may be discerned in, for example, prefrontal cortex (vmPFC) which are of importance. More relevant is the way the signals are computed: which brain regions are involved? What is the nature of their signals? How do they correlate with behavior?

p0350 The potential difference between true and revealed preferences is, in principle, not important for predicting choice, yet it is of utmost important for welfare. At a mundane level, it should be asked whether bankers and brokers ought to engage in eliciting the right prospect-theoretic utility index from their customers through, for example, questionnaires on past or hypothetical choices, to be used to determine optimal portfolios. This only makes sense if prospect theory captures true preferences. If it does not, what is the point of deriving optimal portfolio choice from prospect-theoretic preference profiles (Gomes, 2005)?

p0360 In fact, the mere difficulty of computing optimal choices should cast doubt on the interpretation of prospect theory as a preference theory. Indeed, an organism that truly behaves in a prospect-theoretic way must be endowed with tremendous capacity to compute optimal choices. To put it more bluntly, perhaps organisms would be better off with logarithmic preferences, for which choices are easily computed – for example, one can be myopic (Hakansson, 1971). In addition, this has the advantage that choices maximize survival probabilities (Blume and Easley, 1992)!

p0370 Data in De Martino *et al*. (2006) could be re-interpreted to mean that choice is the result of valuing the appropriateness of some default action, given the stimuli at hand, and, if inappropriate, effort to overcome one's tendency to implement the default action. Recent single-unit recording of neurons in the monkey brain indicates that caudate plays a crucial role in the valuation of default actions. Figure 23.3 provides evidence.

p0380 In the task that generated Figure 23.3, a monkey had to fixate on the center of the screen and at a certain time move its eyes (execute a "saccade") in the direction indicated by a prior stimulus. Only one direction was rewarded, however. To be precise, the chance that one direction would be rewarded was an order of magnitude higher than for the other direction. As it turns out, certain caudate neurons increase their firing rate at trial onset, even before the stimulus, and hence the required direction of motion of the eyes, as well as the amount of the reward, is known. This firing reflects valuation of the default action – namely, to move the eyes in the direction that is most likely to be rewarded. When the stimulus appears, neuronal firing either increases, reflecting confirmation that the default action is correct, or decreases, signaling that the default action is inappropriate (Figure 23.3). In fact, if the percentage of correct saccades is an indication, effort is involved in avoiding the default action. Indeed, Lauwereyns *et al*. (2002) reported that the monkey made more mistakes (i.e., moved its eyes in the wrong direction) when the stimulus required a saccade in the less rewarding direction.

The firing pattern of the caudate neurons also casts doubt on random-utility models (McFadden, 1974) as an explanation of behavior that is at odds with maximization of a fixed utility index. The pattern suggests that the apparent "erratic" behavior of the monkey (its mistakes) is the result of its inability to overcome the default action – for example, if the monkey moved its eyes in the direction that was rewarded with higher probability while the stimulus instructed it otherwise, it was not that its risk tolerance had increased so much as that it wanted to take the chance that its visual p0390

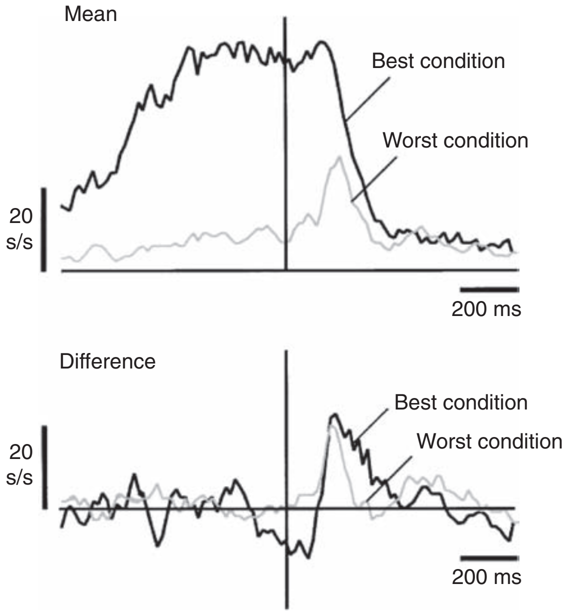


FIGURE 23.3 Average firing rates of a cluster of caudate neurons in the monkey brain that encode the value of eye movements (saccades) in one "preferred" direction. In trials where this direction generates higher reward on average ("Best condition"), increase in firing starts several hundred milliseconds (ms) before stimulus emerges (vertical line) that indicates which direction will be rewarded (top graph, labeled "mean"). Anticipatory increases in firing rates are absent in trials where the "preferred" direction receives the lower reward on average ("Worst condition"). After stimulus presentation, an increase is recorded in average firing rates across trials where the stimulus indicates reward for a saccade in the "preferred" direction, relative to average firing rates when saccades are instructed in the opposite direction. The difference (bottom graph) is the same for both "Best condition" and "Worst condition" trials. From Lauwereyns *et al.* (2002). f0030

inference (of the instructions revealed by the stimulus) was wrong. The mistake was caused by the monkey's inability to overcome its default action (a saccade in the direction of reward with higher probability). Conversely, if the monkey did not make a mistake and moved its eyes in the direction of lower probability of reward, as instructed through the stimulus, it was not that the monkey became less risk-tolerant and did not want to take chances; instead, it overcame its natural inclination to saccade in the direction that tended to be rewarded with higher probability.

p0400 Even if the random-utility model (Box 23.2) and the default-action valuation model predict choice equally well, economists should be aware that neurobiological evidence favors the latter. This is because the welfare implications of the two models are diametrically opposed. According to the random utility model, choice is always optimal. The agent makes different choices in two instances of the same situation only because utility fluctuates. In the default-action valuation model, however, choice is often sub-optimal – for example, the monkey wanted to move its eyes in the opposite direction, but could not because of the effort required to overcome its habit. Now imagine that some external correcting device can be provided using an effort that is less than that required for the monkey personally to overcome its action bias. In that case, welfare will improve. See Bernheim and Rangel (2008) for further discussion of welfare assessment in a context where one needs to distinguish between "true" and "decision" utilities.

b0020 BOX 23.2

RANDOM-UTILITY MODELS

p0070 Random utility models form a class of statistical models widely used to describe actually observed choice behavior. Following standard revealed-preference principles, individual choices are assumed to maximize some underlying utility. The latter is, however, imperfectly observable. Utility, therefore, is split into a modeled part (a function of observable variables, common to all individuals up to a parameter vector) and an orthogonal, random error term (often referred to as "unobserved heterogeneity").

p0080 Specifically, the utility for alternative a is written as

$$U_a = V_a + \varepsilon_a,$$

where V_a is the modeled part of the utility, and ε_a the error term, capturing the uncertainty over the true U_a. The probability of observing alternative a being chosen in the choice set C is therefore

p0090

$$\Pr_C(Y = a) = \Pr(U_a = \max_{c \in C} U_c)$$

p0100 The uncertainty ε_a is frequently modeled as a normal or logistic distribution, yielding probit or logit models, respectively.

VALUATION UNDER PURE RISK: TRADING OFF RISK AGAINST REWARD s0030

p0410 Pure risk is a situation where probabilities are known, either because they are given or because the agent has gone through lengthy training. Economists model choice under pure risk as if the agent maximizes a utility index based on a separation of probabilities (of all possible states of nature) and utilities (of rewards in a state). This separation is known as *probabilistic sophistication* (Marinacci, 2002), and is common to virtually all preference models under pure risk (including prospect theory).

p0420 There appears, however, to be little neurobiological evidence for value computation based on separation of (state) probabilities and (reward) utilities. Instead, brain activation data suggest a separation of encoding of expected reward and of risk.

p0430 Evidence of encoding of expected reward, and the crucial role of the dopaminergic system, goes back a long way, to the investigation of firing of dopaminergic neurons in the monkey midbrain in the face of stochastic rewards (Hollerman and Schultz, 1998). Closer study of the activation of these neurons has revealed that the signals actually reflect *errors* of predicting reward, and these errors form the crucial component of a simple yet versatile and powerful algorithm to learn expected rewards based on the Rescorla-Wagner reinforcement learning rule and extendible to complex multiple stimuli-reward situations (Montague *et al.*, 1996; McClure *et al.*, 2003; O'Doherty *et al.*, 2003). The prediction error equals the difference between actual or updated and (previously) expected (sums of discounted) rewards. A number of chapters in this volume provide details, so we shall not elaborate here (see Chapters 21, 22, 24, and 26).

p0440 One crucial element is missing in this account of (expected) reward learning, namely, risk. Specifically, to learn at the optimal rate, it is important that the decision-maker assesses the risk of making a prediction error. The reasoning is simple. If expecting to make large prediction errors (i.e., if the risk is expected to be high), then one should not change one's prediction much based on a large actual prediction error. In other words, the learning rate should be low. (A few neuroscientists have started to incorporate risk in reinforcement learning, referring to the (prediction) risk as *expected uncertainty* (Yu and Dayan, 2003).For a rigorous treatment, see Preuschoff and Bossaerts, 2007.)

p0450 There are (at least) two ways in which risk assessment could affect learning. One is directly through the learning rate; another is through adjustment of the prediction error. Indeed, a simple scaling of prediction errors by their expected size accomplishes the same (see Preuschoff and Bossaerts, 2007, for details). Evidence for such a type of "adaptive encoding" was recently discovered in firing of dopamine neurons in the monkey brain. Figure 23.4 shows the average firing rates and single-trial firings of dopamine neurons for three types of trials, in the order (top to bottom) of increasing expected size of the prediction error. When reward occurs, the prediction errors are insensitive to the level of anticipated risk, demonstrating that they are somehow "scaled." This finding is further explored in Chapter 21.

p0460 (Expected) reward learning is complex. Not only will it be affected by the expected size of the prediction error (risk); other considerations also play a role, such as how much an optimal predictor might co-vary with the prediction error. For instance, in situations where the environment changes rapidly, past prediction errors become obsolete fast, and hence prediction should rely more on recent prediction errors. Effectively, this means that the learning rate should increase. This intuition has a rigorous underpinning (see Preuschoff and Bossaerts, 2007). Behrens *et al.* (2007) recently provided evidence that humans do adjust their learning rate to the degree of stability.

p0470 In addition to being necessary for optimal learning, risk-encoding could also play a role in the decision-maker's risk *attitude*. Risk-encoding is actually *presumed* to occur in risk-sensitive organisms. How else would their behavior be affected by risk? Not surprisingly, the evidence of risk-encoding in the human and non-human primate brain is overwhelming. Regions where activation appears to be sensitive to risk include the insula, anterior cingulated cortex (ACC) and inferior frontal gyrus (IFG) (see, for example, Critchley *et al.*, 2001; Paulus *et al.*, 2003; Huettel *et al.*, 2005). Some of these regions, such as insula, seem to encode risk *exclusively* (Preuschoff *et al.*, 2006, 2008).

p0480 Moreover, in simple monetary gambles, activation seems to reflect *variance* (or its square root, standard deviation), i.e., the expectation of the squared prediction errors (Preuschoff *et al.* 2006, 2008). As such, the encoding provides a key parameter with which to compute the utility of a risk-averse agent. Indeed, a Taylor series expansion of an expected utility index demonstrates that risk is to be measured in terms of reward variance; all other aspects of risk (skewness, kurtosis, etc.) are secondary. For discussion of applications to finance, see Bali *et al.* (2008), for neuroimaging data on possible encoding of skewness and kurtosis, see Hsu (2006).

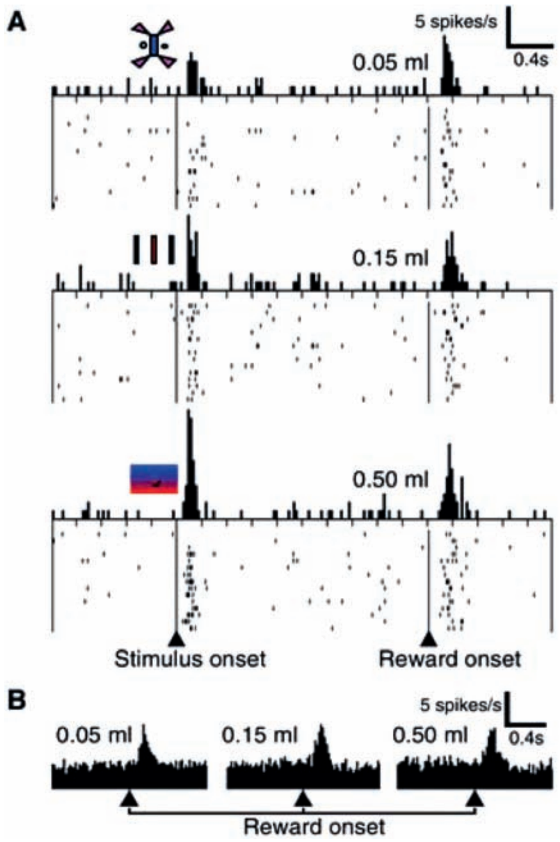


FIGURE 23.4 Single-trial firing rates and corresponding histogram for a single dopamine neuron in ventral tegmental area (a) and histogram of group firing rates (b) in imperative trials where a monkey is presented with stimuli that indicate the size of a random (50% chance) juice reward; results are shown only for rewarded trials. Firing at stimulus onset reflects the size of the expected juice reward. Firing at reward onset reflects the prediction error (positive, since only rewarded trials are retained for the analysis). The firing at reward delivery does not, however, reflect the size of the prediction error, which increases with the size of the reward (e.g., 0.50 ml against an expectation of 0.25 ml in the third case, or 0.05 ml against an expectation of 0.025 ml in the first case). Firing at reward delivery reflects a scaled prediction error: the prediction error divided by the anticipated size of the prediction error (prediction risk). From Tobler *et al.* (2005). f0040

p0490 As with expected reward signals in the dopaminergic system, activation correlating with risk in some regions actually reflects risk-prediction *errors* – i.e., the difference between the square-size of the prediction error and its expectation (the variance). Specifically, phasic activation in anterior insula exhibits strong correlation with risk-prediction errors (Preuschoff *et al.*, 2008). Little is known, however, about the precise neurobiological architecture and algorithms of risk learning. We only know that risk prediction and risk-prediction errors are encoded.

p0500 One may object that a risk-prediction error can be recorded in the paradigm of Preuschoff *et al.* (2008). Indeed, the risk-prediction error in a trial is irrelevant to making better forecasts in future trials, as the trials are independent. That is, there is no relevant learning. There are three reactions to this objection. First, this finding is not unique to risk-prediction errors. Reward-prediction errors can similarly be recorded in situations where learning is irrelevant because trials are independent (see, for example, Fiorillo *et al.*, 2003; Preuschoff *et al.*, 2006). Second, encoding of prediction errors is relevant for cases where learning unexpectedly *does* matter. For example, if an experimenter tells a subject that trials are independent (as in Preuschoff *et al.*, 2006) yet is lying (unlike in Preuschoff *et al.*, 2006), then if the brain has prediction errors at hand it can manage the situation better. (In fact, this is one of the prime reasons why deception is frowned upon in economic experiments, but it would not be a compelling one if people always blindly believed the experiment.) In this sense, the emergence of prediction errors in situations where learning is irrelevant in principle is equivalent to the presence of valuation signals in imperative trials.

p0510 Third, prediction-error encoding does sometimes have behavioral implications. Choices across independent trials sometimes reflect belief updating as if outcomes are actually dependent. Such belief updating presumes encoding of prediction errors. A classical example is underweighting of small probability events in experiments with trial-by-trial reward feedback (see Hertwig *et al.*, 2004).

p0520 Finally, risk-encoding may play a role in learning, in addition to guiding choice. For learning to be effective, however, risk-encoding should be *objective*. In particular, risk aversion or risk tolerance should not affect risk perception. Imagine that more risk-averse agents overestimate risk. Since perceived risk decreases the learning rate, this would mean that risk-averse agents learned more slowly. That would give them a handicap in competition with agents who are more risk tolerant. To date, it is not known to what extent learning is affected by risk aversion, or to what extent the brain encodes a dual-risk signal – one subjective, to guide choice, and one objective, to drive learning.

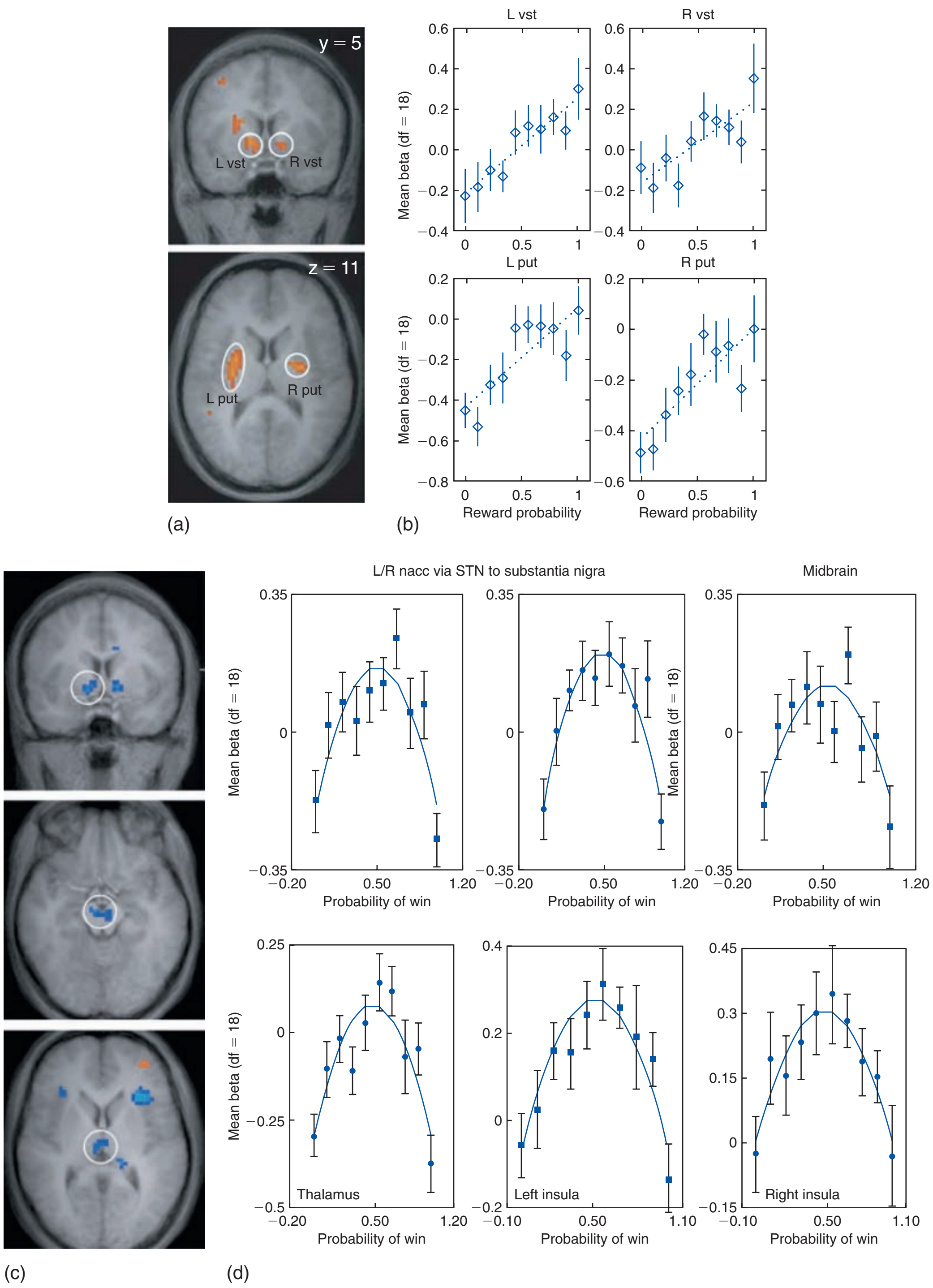
p0530 The picture that emerges from the extant findings is one where the brain values risky gambles by evaluating their expected reward and risk separately (Figure 23.5). The separate evaluations are then merged to generate a total valuation signal, detectable in (for example) prefrontal cortex (PFC). Figure 23.6 shows how PFC activation increases with expected reward for all subjects, but decreases with risk for risk-averse subjects and increases with risk for risk-seeking subjects.

p0540 To date, it is not known how the various signals related to risk and expected reward in disparate brain regions are integrated to deliver the aforementioned total value signal in PFC. A recent study, however, used canonical correlations analysis to extract a value index from joint activation of insula, ventral striatum, and putamen, exploiting the recorded correlation in baseline activations across these regions. The resulting "neuronal" utility index was found to correlate significantly with activation in the very part of PFC identified in Tobler *et al.* (2007; see also Bruguier *et al.*, 2008).

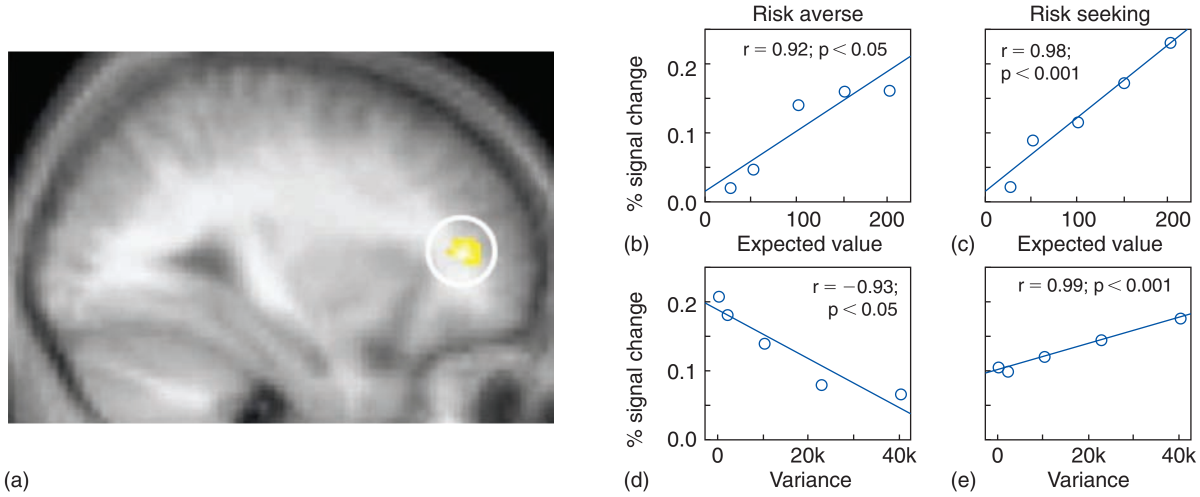
EXTENDING THE REWARD–RISK COMPUTATIONAL MODEL TO AMBIGUITY s0040

p0550 So far, we have considered only pure risk – i.e., situations where probabilities are known. The evidence points towards a computational model whereby the brain computes value by separately encoding expected reward and risk, and combining the results. Such a computational model is known to approximate well to many types of utility functions (Bali *et al.*, 2008), including prospect theory (Agren, 2006).

p0560 In many situations in real life, probabilities are unknown or only partially known. This is called *ambiguity* in decision theory (Ellsberg, 1961). A stylized fact in decision experiments is that many individuals are ambiguity averse (Camerer and Weber, 1992). A number of models exist that account for decisions under ambiguity. Perhaps the best known are the maxmin utility model of Gilboa and Schmeidler (1989), and its extension, the α-maxmin utility model (Ghirardato *et al.*, 2004). In the latter, the worst- and best-case scenarios provide anchor points for beliefs. In one extreme case ($\alpha = 1$), beliefs are entirely determined by the worst-case scenario, and the α-maxmin utility model coincides with that of Gilboa and Schmeidler (1989). In the other extreme case ($\alpha = 0$), beliefs are solely based on the best-case scenario. For $\alpha > 0.5$, the decision-maker is ambiguity averse; for $\alpha < 0.5$, she is ambiguity seeking; at $\alpha = 0.5$, she is ambiguity neutral – i.e. her beliefs are the average of the worst-case and best-case scenarios.



f0050 **FIGURE 23.5** (Immediate) activation in striatal areas (a) increases with the probability of a fixed reward (b) in the human brain in gambles where the probability of reward fluctuates between 0 and 1. As such, activation correlates with expected reward. (Delayed) activations in sub-cortical areas and insula (c) change quadratically with probability of reward (d), with a peak at 50% chance of winning, and minima at certainty of loss or gain. As such, activation correlates with risk, measured as reward variance. Vertical line segments indicate 95% confidence intervals. From Preuschoff *et al*. (2006), and unpublished data.



f0060 **FIGURE 23.6** Activation in medial prefrontal cortex correlates with expected utility of a gamble in imperative trials (a). This activation increases in expected reward both for risk-averse (b) and risk-seeking (c) subjects; risk attitudes are measured through revealed preference in free-choice trials. The same activation decreases with risk of the gamble (measured as reward variance) for risk-averse subjects (d). It increases with risk for risk-seeking subjects (e). Adapted from Tobler *et al*. (2007).

p0570 For instance, the decision maker may bet on the drawing of red, green, and blue balls from an urn. She earns \$1 if the ball drawn is red or green; if the ball is blue, she loses \$1. One-third of the balls are red. The number of green and blue balls is unknown, but the decision-maker knows that at least one-sixth of the balls are green, and one-twelfth are blue. A decision-maker with $\alpha = 0$ will only consider the worst-case scenario, the minimum probability of winning, $p_{min} = 1/2$ $(=(1/3) + (1/6))$; this will fix her beliefs. When $\alpha = 1$, the maximum probability, p_{max}, determines her beliefs; $p_{max} = 11/12$ $(=1 - 1/12)$. In general, the belief that the decision-maker wins equals $\alpha\ p_{min} + (1 - \alpha)\ p_{max}$. Without loss of generality, we can set the utility of winning \$1 equal to 1 and the utility of losing \$1 equal to 0. Hence, the general expression of the expected utility of the decision maker is $\alpha\ p_{min} + (1 - \alpha)\ p_{max}$. See Box 23.3 for further discussion of the α-maxmin model.

p0580 As with expected utility under pure risk, the α-maxmin utility model is equivalent (this time, it is *not* an approximation) to one where utility is computed based on a tradeoff between mean and risk. In this case, the mean is just the average of the minimum and maximum probabilities:

$$\bar{p} = \frac{1}{2} p_{min} + \frac{1}{2} p_{max}.$$

The risk c, in contrast, is determined by the amount ambiguity:

$$c = p_{max} - p_{min}.$$

Utility, then, is some weighted average of β and c:

$$U = \gamma_1 \bar{p} + \gamma_2 c.$$

Simple algebra reveals that:

$$\alpha = \frac{1}{2}\gamma_1 - \gamma_2.$$

$$1 - \alpha = \frac{1}{2}\gamma_1 + \gamma_2.$$

So $\gamma_1 = 1$ and $\gamma_2 = 1/2 - \alpha$.

p0590 Therefore, under both pure risk and ambiguity, there exists an equivalent representation of utilities in terms of a model that trades off mean against risk. As such, it can be conjectured that brain regions involved in encoding mean and risk in situations of pure risk (striatum, anterior cingulate cortex, insula, inferior frontal gyrus, …) are also involved in encoding these parameters when there is ambiguity. As the weights on mean and risk need not be the same, it is to be expected that the intensity of the brain activations differs depending on the situation.

p0600 This is supported by the findings of Hsu *et al*. (2005) and consistent with those of Huettel *et al*. (2006). The former found the striatum to be differentially activated under risk relative to ambiguity, whereas the amygdala and lateral orbitofrontal cortex showed the reverse. Furthermore, they provide both neuroimaging and brain-lesion data support for the idea that a common neural mechanism underlies both risk and b0030 BOX 23.3

MODELS OF DECISION-MAKING UNDER AMBIGUITY

p0110 A number of models exist that account for decisions under ambiguity. Perhaps the best known is the maxmin utility (MEU) model (Gilboa and Schmeidler, 1989), and its extension the α-maxmin (α-MEU) model (Ghirardato *et al*., 2004).

p0120 The former replaces the classic independence axiom with a weakened version (certainty-independence), thereby yielding the utility representation

$$MEU(f) = \min_{P \in C} \int_S u(f)dP,$$

p0130 where C is the set of probability measures on the set of possible states S. Under MEU, the decision maker considers only the worst-case scenario.

p0140 The α-maxmin model generalizes MEU by allowing the decision-maker to consider a mixture of the worst and best-case scenarios, which provide anchor points for beliefs. That is,

$$\alpha MEU(f) = \alpha \min_{P \in C} \int_S u(f)dP + (1-\alpha)\max_{P \in C} \int_S u(f)dP.$$

p0150 When $\alpha = 1$, beliefs are entirely determined by the worst-case scenario, and the αMEU utility model coincides with MEU. In the other extreme case ($\alpha = 0$), beliefs are solely based on the best-case scenario. For $\alpha > 0.5$, the decision-maker is ambiguity averse; for $\alpha < 0.5$, she is ambiguity loving; at $\alpha = 0.5$, she is ambiguity neutral; her beliefs are the average of the worst-case and best-case scenarios.

ambiguity. First, the striatal activity was significantly correlated with the expected value of subjects' choices in both the risk and ambiguity conditions (Hsu *et al*., 2005; see also Figure 23.7). Even stronger support is provided by the fact that patients with lesions to the lateral orbitofrontal cortex, as opposed to a comparison group of temporal-lobe lesion patients, were both risk *and* ambiguity neutral. Therefore, the difference between risk and ambiguity at the level of the reward system appears to be one of degree.

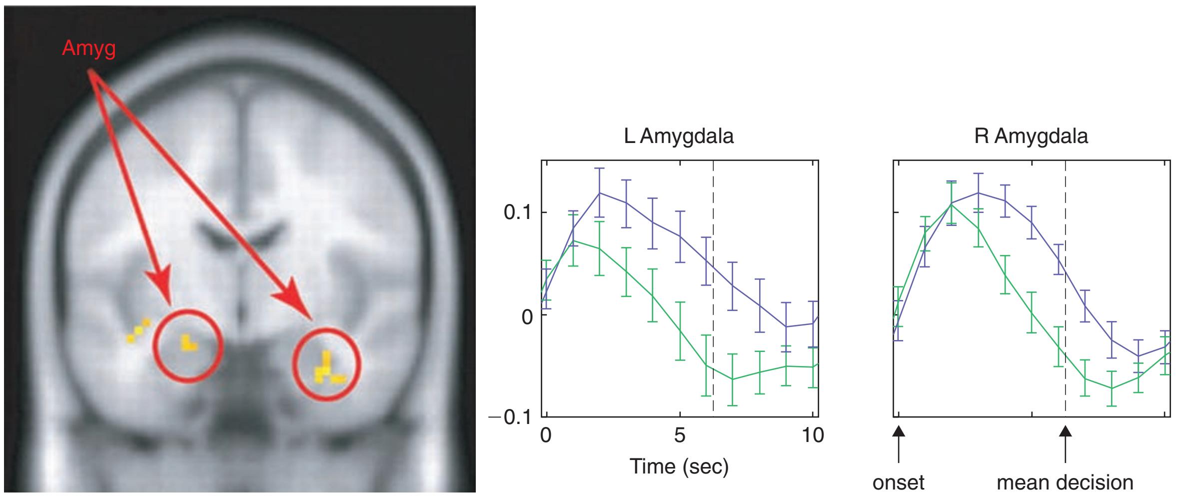
p0610 Closer inspection reveals that the mean-risk model is actually more general than the α-maxmin model; the latter imposes restrictions on the weights and on the mean-risk model. In fact, the mean-risk mode could be derived as a hierarchical Bayesian updating model, whereby the decision-maker is first agnostic about the true probability of winning (in the example above it is uniform between 1/3 and 1), but then uses the midpoint (the average between minimum and maximum probabilities) as a cue to where the true probability lies. Specifically, the midpoint is assumed to be drawn from a uniform distribution centered on the true probability and with a range equal to the difference between the maximum and minimum probabilities. The posterior mean and variance of the true probability are given by $\bar{p}$ and c, respectively.

p0620 The connection with updating and learning is important and illuminating in light of the findings of amygdala activation in Hsu *et al*. (2005). Significant activation of amygdala is rare or non-existent under pure risk. When contrasting choices involving ambiguity against those involving only pure risk, however, amygdala activation is strong. This could be consistent with activation of amygdale in, for example, novelty tasks (Rutishauser *et al*., 2006), and suggests that the amygdala signals a need for learning of the unknown probabilities. Further research is needed to discover the true role of amygdala under ambiguity.

p0630 More generally, further experiments are needed to verify the extent to which valuation in the context of ambiguity is based on a computational model that trades off mean against risk. These experiments should extend previous studies pairing ambiguous with pure-risk (or zero-risk) gambles (Hsu *et al*., 2005; Huettel *et al*., 2006), and focus on choice between gambles with differing levels of ambiguity.

EXTENDING THE REWARD–RISK COMPUTATIONAL MODEL TO DECISIONS INVOLVING EQUITY AND EFFICIENCY s0050

p0640 The separation between expected reward and variance we observe in neural valuation of risky gambles



(a) (b)

f0070 FIGURE 23.7 Activation in amygdala is higher in free-choice trials involving an ambiguous option than those only involving pure risk. (a) Location of activation. (b) fMRI activation as a function of time since trial onset; dashed vertical line indicates mean time of choice; blue line is average activation in trials with ambiguity; green line is average activation in trials with pure risk; vertical line segments indicate 95% confidence intervals. Autocorrelation in time series is meaningless, as is scale of activation; only the difference in activation matters. From Hsu *et al.* (2005).

may also carry over to choices that require tradeoffs between efficiency and inequity. The latter is a typical problem in studies of distributive justice. A deep theoretical connection exists between decision-making under uncertainty and the measurement of inequity (Atkinson, 1970; see also Box 23.4). The central concern in both is the comparison of frequency distributions *f*(*y*). In decisions under uncertainty, *f*(*y*) denotes the probability of state *y* occurring, whereas in the case of inequity, *f*(*y*) is the frequency distribution on income or other variables of interest for person *y*.

b0040

BOX 23.4

MEASUREMENT OF INEQUITY AND DECISION-MAKING UNDER RISK

p0160 Central in the connection between measurement of inequity and decision-making under risk is the idea of welfare. This is uncontroversial in the latter, and is traditionally meant to denote individual welfare, or utility. As pointed out in Dalton (1920), however, judgments of social welfare underlie the conception of any inequity measure. If we assume that the social welfare function is additively separable and symmetric in income, we arrive at the following:

$$W = \int_0^{\bar{y}} U(y) f(y) dy.$$

p0170 This form is immediately familiar in its resemblance to the standard expected utility representation. In fact, many of the concepts in decision-making under risk, e.g., second-order stochastic dominance, mean-preserving spread, have formally identical counterparts in measurements of inequity, e.g., Lorenz dominance, principle of transfer, respectively (Atkinson, 1970). The assumptions of additive separability and symmetry, interestingly, can in addition be derived axiomatically via an appeal to decision-making under risk (Harsanyi, 1978).

p0650 In experiments involving efficiency and inequity tradeoffs in a distributive justice task, Hsu *et al.* (2008) found a separation of regions encoding efficiency and inequity in much the same way as that for expected reward and variance. Specifically, efficiency is encoded in striatal regions including the caudate and putamen, whereas inequity is encoded in the insular cortex. This utilization of similar neural mechanisms for what on the surface are very different decisions provides a strong indication that these are very general mechanisms for the computation of decisions in a variety of different situations.

Acknowledgments s0060

Financial support from the Swiss Finance Institute and the Swiss National Science Foundation (Finrisk NCCR) is gratefully acknowledged. p0660

References

Agren, M. (2006). Prospect theory and higher moments. Working Paper, Uppsala University.

Atkinson, A.B. (1970). On the measurement of inequality. *J. Econ. Theory* 2, 244–263.

Bali, T.G., Demirtas, K.O., Levy, H. *et al*. (2008). Is there a relation between downside risk and expected stock returns? *J. Financial Quant. Anal.*, (forthcoming).

Bechara, A. and Damasio, A.R. (2005). The somatic marker hypothesis: a neural theory of economic decision. *Games Econ. Behav.* 52, 336–372.

Bechara, A., Damasio, H., Tranel, D. *et al*. (1997). Deciding advantageously before knowing the advantageous strategy. *Games Econ. Behav.* 275, 1293–1295.

Bechara, A., Damasio, H., and Damasio, A.R. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *J. Neurosci.* 19, 5473–5481.

Behrens, T.E., Woolrich, M.W., Walton, M.E. *et al*. (2007). Learning the value of information in an uncertain world. *Nat. Neurosci.* 10, 1214–1221.

Bernheim, B. D. and A. Rangel (2008). Beyond revealed preference: toward choice-theoretic foundations for behavioral welfare economics. National Bureau of Economic Research Working Paper Series No. 13737.

Berns, G.S., Capra, M., Chappelow, J. *et al*. (2007). Nonlinear neurobiological probability weighting functions for aversive outcomes. *NeuroImage* 9, 2047–2057.

Blume, L. and Easley, D. (1992). Evolution and market behavior. *J. Econ. Theory* 107, 95–130.

Bruguier, A., Preuschoff, K., Quartz, S. *et al*. (2008). Investigating signal integration with canonical correlation analysis of fMRI brain activation data. *NeuroImage* 41, 35–44.

Camerer, C. and Weber, M. (1992). Recent developments in modeling preferences – uncertainty and ambiguity. *J. Risk Uncertainty* 5, 325–370.

Critchley, H., Mathias, C., and Dolan, R. (2001). Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron* 29, 537.

Dalton, H. (1920). The measurement of the inequality of incomes. *Economic J.* 30, 348–361.

De Martino, B., Kumaran, D., Seymour, B. *et al*. (2006). Frames, biases, and rational decision-making in the human brain. *Science* 313, 684–687.

Diaconis, P. and Freedman, D. (1986). On the consistency of Bayes estimates. *The Annals of Statistics* 14(1), 1–26.

Dunn, B.D., Dalgleish, T., and Lawrence, A.D. (2006). The somatic marker hypothesis: a critical evaluation. *Neurosc. Biobehav. Rev.* 30, 239–271.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *Q. J. Economics* 75, 643–669.

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

Ghirardato, P., Maccheroni, F., and Marinacci, M. (2004). Differentiating ambiguity and ambiguity attitude. *J. Econ. Theory* 118, 133–173.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *J. Math. Econ.* 18, 141–153.

Gomes, F.J. (2005). Portfolio choice and trading volume with loss-averse investors. *J. Business* 78, 675–706.

Hakansson, N.H. (1971). Multi-period mean-variance analysis: toward a general theory of portfolio choice. *J. Finance* 26, 857–884.

Harsanyi, J.C. (1978). Bayesian decision theory and utilitarian ethics. *Am. Econ. Rev.* 68, 223–228.

Hertwig, R., Barron, G., Weber, E.U. *et al*. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Sci.* 15, 534–539.

Hollerman, J.R. and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* 1, 304–309.

Hsu, M. (2006). Three correlated essays on the neural foundations of economic decision making. PhD thesis, Faculty of Social Sciences, Pasadena, California Institute of Technology.

Hsu, M., Bhatt, M., Adolphs, R. *et al*. (2005). Neural systems responding to degrees of uncertainty in human decision making. *Science* 310, 1680–1683.

Hsu, M., Anen, C., and Quartz, S. (2008). The right and the good: distributive justice and neural encoding of equity and efficiency. *Science*, (in press)..

Huettel, S., Song, A., and McCarthy, G. (2005). Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. *J. Neurosci.* 25, 3304–3311.

Huettel, S.A., Stowe, C.J., Gordon, E.M. *et al*. (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron* 49, 765–775.

Kahneman, D. and Tversky, A. (1992). Advances in prospect theory: cumulative representation of uncertainty. *J. Risk Uncertainty* 5, 297–324.

Kahneman, D., Wakker, P.P., and Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *Q. J. Economics* 112, 375–405.

Lauwereyns, J., Takikawa, Y., Kawagoe, R. *et al*. (2002). Feature-based anticipation of cues that predict reward in monkey caudate nucleus. *Neuron* 33, 463–473.

Lucas, R. (1976). Econometric policy evaluation: a critique. *Carnegie-Rochester Conference Series on Public Policy* 1, 19–46.

Marinacci, M. (2002). Probabilistic sophistication and multiple priors. *Econometrica* 70, 755–764.

Markowitz, H. (1952). Portfolio selection. *J. Finance* 7, 77–91.

McClure, S.M., Berns, G.S., and Montague, P.R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38, 339–346.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In: P. Zaremba (ed.), *Frontiers in Econometrics*. New York, NY: Academic Press, pp. 105–142.

Montague, P.R., Dayan, P., and Schultz, W. (1996). A framework for mesencephalic dopamine systems based on predictive hebbian learning. *J. Neurosci.* 16, 1936–1947.

O'Doherty, J.P., Dayan, P., Friston, K. *et al*. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron* 38, 329–337.

Paulus, M.P., Rogalsky, C., Simmons, A. *et al*. (2003). Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *NeuroImage* 19, 1439–1448.

Preuschoff, K. and Bossaerts, P. (2007). Adding prediction risk to the theory of reward learning. *Ann. NY Acad. Sci.* 1104, 135–146.

Preuschoff, K., Bossaerts, P., and Quartz, S.R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* 51, 381–390.

Preuschoff, K., Quartz, S.R., and Bossaerts, P. (2008). Human insula activation reflects risk prediction errors as well as risk. *J. Neurosci.* 28, 2745–2752.

Rutishauser, U., Mamelak, A.N., and Schuman, E.A. (2006). Single-trial learning of novel stimuli by individual neurons of the human hippocampus–amygdala complex. *Neuron* 49, 805–813.

Segal, U. (1987). The Ellsberg Paradox and risk-aversion – an anticipated utility approach. *Intl Econ. Rev.* 28, 175–202.

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

Tobler, P.N., O'Doherty, J.P., Dolan, R.J. *et al*. (2007). Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *J. Neurophysiol.* 97, 1621–1632.

Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* 19, 181–189, (1 %R doi:10.1111/j.1460–9568.2004.03095.x).

Yu, A.J. and Dayan, P. (2003). Expected and unexpected uncertainty: ACh and NE in the neocortex. *Adv. Neural Inf. Proc. Syst.* 15, 157–164.