

c00030

# The Trouble with Choice: Studying Decision Variables in the Brain

Greg S. Corrado, Leo P. Sugrue, Julian R. Brown and  
William T. Newsome

## O U T L I N E

s0010	<b>Introduction</b>	<b>461</b>	<b>Validating Behavioral Performance</b>	<b>472</b>	s0140
s0020	<b>Direct Methods for Accessing Decision Variables</b>	<b>463</b>	<i>Simple Predictive Performance</i>	472	s0150
s0030	<i>Direct Self-report</i>	463	<i>Generative Performance</i>	474	s0170
s0040	<i>Incentive Compatible Mechanism Design</i>	464	<b>The Quest for Neural Correlates</b>	<b>474</b>	s0180
s0050	<i>Simple Revealed Preference</i>	464	<i>Matching Behavior and Value-based Action</i>		s0190
s0060	<b>Model-based Approaches</b>	<b>465</b>	<i>Selection in Area LIP</i>	475	
s0070	<i>Styles of Modeling</i>	466	<i>fMRI Correlates of Reinforcement Learning Models</i>	475	s0200
s0100	<i>Specifying the Details of the Model</i>	467	<b>Conclusion</b>	<b>476</b>	s0210
s0130	<i>Parameter Estimation</i>	471	<i>References</i>	477	s0220

## INTRODUCTION

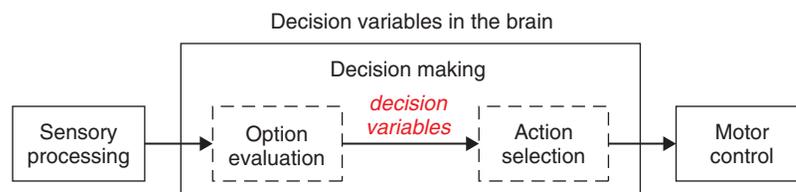
s0010

p0010

Our lives are defined by the decisions we make. From simple acts like picking an item from a lunch menu, to complex choices such as selecting a partner or judging someone guilty of a crime, our decisions vary hugely in complexity and consequence. Given the myriad external and internal factors that influence our choices, it might seem surprising that the underlying machinery of decision-making is amenable to scientific analysis at all. Yet in recent years psychologists, economists, and neuroscientists have made significant inroads into this intriguing but mysterious domain. The inquiry has advanced farthest in the arena of sensory-based decision-making, where subjects are

asked to judge an objective quality of an external stimulus (reviewed in Romo and Salinas, 2003, and Gold and Shadlen, 2007). In this chapter, however, we will focus on the more elusive area of “value-based” decision-making, which is a topic of primary interest in neuroeconomics (Sugrue *et al.*, 2005).

It is useful to start with three definitions. First, for our purposes, a *decision* occurs when an organism, confronted by several discrete options, evaluates the merits of each and selects one to pursue. In the laboratory, we study decisions by placing subjects in a controlled environment and offering them two or more mutually exclusive options. Subjects communicate their choices by making overt motor responses such as button presses or eye movements, motivated by the



f0010 **FIGURE 30.1** Decision variables in the brain. In an (over) simplified feed-forward model of the brain, sensory systems provide data to decision-making circuits, which in turn direct motor systems. The machinery of decision-making can be further subdivided into mechanisms that evaluate options and those that subsequently select actions. Decision variables can then be operationally defined as the currency of communication between these two sub-stages.

promise of reward (or the threat of punishment) contingent on the responses they make.

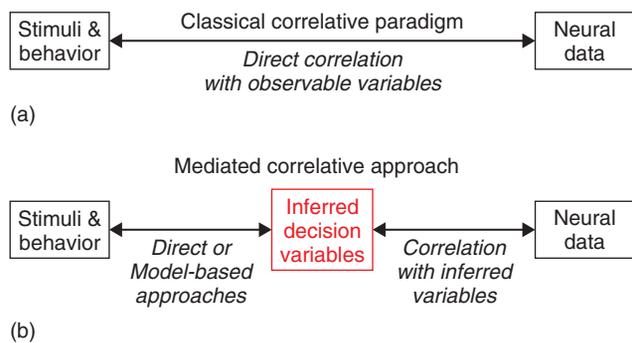
p0030 Second, a *value-based* decision is one in which the organism's choice is not mandated by the immediate sensory characteristics of the options (as when reading an eye chart at the doctor's office), but rather by the subjective experience and preference of the individual (as when selecting a fine wine in a restaurant). Typically, there are no unambiguously "correct" answers in value-based decisions, which is where the trouble begins from a scientific perspective. Unlike simple sensory judgments, choices rendered in value-based tasks depend on unseen factors internal to the decision-maker, and it is here that our third definition comes into play. (Occasionally, of course, there are sensory judgments that lack an objectively "correct" answer, such as the famed Necker Cube, or value-based decisions that have an obviously "correct" answer under only very weak assumptions – for example, would you rather have a 50% chance of winning \$10 or a 50% chance of losing \$10? – but these cases are rare. Typically, sensory judgments hinge on discovering some real objective state of the external world, and value judgments depend on the assessment of outcomes whose worth can be genuinely contentious and idiosyncratic to the individual decider. Naturally, most real-world decisions require both sensory and value-based assessments: what is the state of the world, and how desirable is that state?)

p0040 *Decision variables* are quantities internal to the subject's decision process that summarize properties of the available behavioral options relevant to guiding choice. As illustrated in Figure 30.1, decision variables can be thought of as linking the processes of option evaluation and action selection. These variables are by their very nature subjective, expressing the decider's private estimation of the attractiveness of the available options. To the outside observer such internal states are necessarily hypothetical – indeed, behaviorists might go so far as to call them fictitious. In contrast, neuroeconomists, guided by the conviction that decisions are mediated by scientifically accessible processes within the brain, embrace the study of internal decision variables as a

central goal. We would like to understand where and how these variables are computed within the brain, and how they subsequently guide action selection. Our success or failure in this endeavor will ultimately be judged by our ability to "read out" the state of a decision variable directly from measurements of brain activity, and by our ability to produce predictable changes in choice behavior by manipulating decision variables at the neural level. Before reaching either of these goals, however, neuroeconomists must first confront and solve two related problems.

First, we must find a way to deal with the hypothetical nature of decision variables. We have no *a priori* way of knowing what decision variables guide particular choices. Even in very simple value-based decision-making situations, the factors guiding choices may include any property of a behavioral option that might impact its desirability: reward magnitude, reward quality, risk, uncertainty, delay, cost, etc. While each of these elements might individually constitute a decision variable in its raw form, they might also be transformed or combined in some complicated way (as by the economist's utility function). Moreover, any of these variables – raw, transformed, or combined – might be encoded within the brain in absolute terms, relative to other available options, or even relative to the subject's expectations. In yet more complex situations, decision variables might also include additional hidden factors ranging from the physiological state of the subject (e.g. hunger) to subtle social pressures to make the "right" choice.

The second problem follows directly from the first: because we do not know *a priori* what decision variables are actually used by the brain, the standard correlative methods employed by systems neuroscientists are seriously compromised. Historically, success in systems neuroscience has come from manipulating a variable of interest (i.e., a sensory stimulus or a motor act), comparing neural activity across experimental conditions, and attributing observed differences to functional processing within the central nervous system (Figure 30.2a). Note that the "variable of interest" might be something as simple as the pitch of a



**FIGURE 30.2** Topology of correlative methods. (a) The classical correlative approach looks for direct correlations between observable variables, such as the presented sensory stimuli or the subject's overt behavior, and physiological signals, such as neural firing rates or blood oxygen-level dependent (BOLD) magnetic resonance signals (fMRI). (b) In the study of decision-making, it is generally necessary to use a mediated correlative approach – inferring abstract decision variables which depend on stimuli and predict behavior, and in turn correlating these computed variables with physiological measurements.

presented auditory tone (e.g., Reale and Imig, 1980), or something as subtle as the attribution of emotional states to self versus other (e.g. Ochsner *et al.*, 2004). Similarly, the “measure of neural activity” might be any available measure of brain physiology: the delivery of oxygen to a particular brain region measured via fMRI, the firing of action potentials by individual neurons within that area measured through electrophysiology, or even the strength of synaptic connectivity between pair of neurons measured by future technologies. Studying decision variables with this approach becomes difficult when we do not already know what decision variables operate in a given situation.

We can try to solve these problems by “bootstrapping” our way into the decision-making system. To do this we make a reasonable guess about the decision variables operating in a given task, and then try to identify correlates of these variables in the brain (Figure 30.2b). The neural signals that we identify through this process may help us reject certain candidate decision variables and form new hypotheses about others. However, identifying meaningful signals in the first place hinges on the accuracy of our initial guess at the candidate decision variables. To improve the accuracy of this guess, we must begin with what *is* observable – the subject's choices. Through careful analysis of behavior, we hope to form good approximations to the hidden decision variables.

Because this process is so crucial, we devote this chapter to a review of the most promising approaches to recovering hidden decision variables from behavior. First, we consider “direct” methods of inferring decision variables from patterns of choice. We then consider “model-based” approaches that construct an

explicit model of the entire decision-making process, before going on to describe approaches to validating candidate decision variables by rigorously comparing their predictions to actual behavior. Finally, we discuss some recent studies in which techniques for estimating decision variables have been successfully applied to search for neural correlates.

## DIRECT METHODS FOR ACCESSING DECISION VARIABLES

Several techniques are available for directly inferring subjective decision variables from behavior. These “model-free” methods are simple in the sense that they do not require mathematical descriptions of how decision variables are computed, although some “model-free” techniques do leverage assumptions about how decision variables guide action selection or how decision variables are combined. These conditions, though, are more modest than the detailed mathematical models of decision-making we consider later in the chapter. However, sometimes this simplicity comes at the cost of cumbersome or fragile experimental design. To begin with, we consider the most obvious approach to accessing decisions variables: self-report.

### Direct Self-report

No approach to accessing internal states could be more straightforward than asking subjects to simply report their feelings. For example, to assess how a subject values a particular option, we could present the option and ask the subject to rate its value on a linear scale. Simple though it is, direct self-report is rarely used as a primary technique for accessing decision variables, for a number of reasons. Psychologists stress subjects' inability to access or correctly identify the true cause of their actions, and the vulnerability of their answers to the precise manner in which the question is framed (see Ross and Nisbett, 1991). Economists, on the other hand, criticize the “lack of incentive” for subjects to report their beliefs truthfully, in the absence of which subjects might report randomly, or even deliberately deceive. Neurophysiologists, meanwhile, dismiss the approach entirely, citing the almost insuperable difficulties in applying such methods to non-human animals. Regardless of the reason, the consensus is clear: unadorned self-report is too unreliable and limited a method to provide robust metrics of hidden decision variables. Fortunately, there are refinements of self-report techniques that promise greater reliability.

s0040 **Incentive Compatible Mechanism Design**

p0110 To address the economist's primary criticism of self-report methods for assessing value, we can employ the principles of *incentive compatible mechanism design*. (For those unfamiliar with the term, *mechanism design* is the field dedicated to the design of social institutions such as voting systems, auction rules, simple games, and market structures, while *incentive compatibility* is the lofty objective that these institutions should function as the designer intended even if the agents who make up the social group are perfectly selfish and fiendishly clever – i.e. that the “incentives” of the game should render individual selfishness “compatible” with the desired overall social outcome.) The idea itself is reasonably simple: ask subjects to report their preference for a particular option in units we understand (such as dollars), and give them a reason to answer honestly. The mechanism employed is usually some sort of virtual auction, where the “prize” is the option the subject is valuing and the “bid” is the value the subject reports. By participating in this virtual auction subjects communicate their valuation of options, much like direct self-report. In this case, however, it is incentive compatible for subjects to report their valuation truthfully because the details of the auction ensure that if they lie, they will on average be less happy with the final outcome than if they tell the truth (although unfortunately, these detailed rules are sufficiently convoluted that it may be only the experimenter's word that motivates many subjects toward honesty, rather than the direct understanding that honesty is truly the best strategy). Such auction-based incentive compatible self-report mechanisms have been used to extract valuations of items from human subjects that predict subjects' choices in other contexts and correlate with activity in neural circuits implicated in the process of valuation (see, for example, Plassmann *et al.*, 2007).

p0120 While this approach meets the economists' demands for motivating subjects toward honesty, psychologists and neurophysiologists tend to be less satisfied. The psychologist's chief objection is that underlying decision variables are intrinsically difficult to access via introspection, even under proper motivation. To wit, one might wonder if eBay's auction selling prices reflect the bidder's true cold assessment of the value of an item, or are more often than not contaminated by the thrill of competition or clouded by the abstractions of credit cards and PayPal. For the neurophysiologist, meanwhile, who struggles to study neural activity in experimental animals that lack even basic facility with the abstractions of language or currency, any method that relies upon *direct* reports of

decision variables, incentive compatible or not, is of limited practical usefulness.

**Simple Revealed Preference**

Simple *revealed-preference* techniques offer another alternative to self-report, allowing us to infer an option's relative value directly from patterns of choice across pairs of options. At their core, these revealed-preference methods rely on the relatively conservative assumption that when choosing between two options A and B, a subject will choose A more frequently than B if (and only if) A is more desirable:

$$D_A > D_B \leftrightarrow p(\text{choice} = A) > p(\text{choice} = B) \quad (30.1)$$

where we have used  $D$  to denote the total subjective desirability of an option presumed to combine all the subject's preferences over possible outcomes, the probability of outcomes, risk, cost, etc. (Two points should be remembered here. First, the term *revealed preference*, as it is used in economics, refers to a vast body of work produced over the past 70 years, complete with a collection of axioms relating choice patterns to rationality. Here we do not mean to invoke that entire literature, but only Samuelson's original observation that preference could be inferred from choice. Second, in some sense, even the incentive compatible self-report mechanism mentioned above is a form of revealed preference, insofar as casting the task as an auction causes subjects to “reveal” their valuations through their bids rather than directly “reporting” them to the experimenter. Moreover, the model-based methods we consider in the following section leverage revealed preference in that behavior is the medium from which valuations are reconstructed. In this section, we restrict ourselves to what might be called *naked* revealed preference – simple efforts to estimate valuation by directly tabulating choices across pairs of options.)

An example of a simple revealed-preference design is one in which subjects are given a large number of paired choices between  $n$  units of reward X, and  $m$  units of reward Y. By finding the point at which a subject is indifferent to the two options the experimenter can infer the effective exchange rate between X and Y (for example, three units of X are as desirable as two units of Y). Using such *indifference points*, one can obtain an ordinal ranking of the desirability of any amount of either reward quantity, or even combinations of the two. (Assumptions about how values combine must be validated by comparing predicted choice to observed choice, as we detail below in “Validating behavioral performance.”) The ranking of option values

can then be used to search for neural correlates, thereby identifying candidate structures or circuits for the neural representations of subjective desirability. A chief advantage of this approach over the methods discussed so far is that it is readily applicable to the study of value-based decision making in non-human animals. For example, it has been successfully employed in insulating neural decision variables in monkeys choosing between different quantities and flavors of fruit juice (Padoa-Schioppa and Assad, 2006).

p0150 An unfortunate limitation of many such simple revealed-preference designs is their assumption of *stationarity*. Because these methods generally rely on combining data across many trials to compute indifference points, they typically assume that the decision variables summarizing an option are identical every time that option is presented. (It is in principle possible to include the effects of non-stationarities when deducing revealed preferences, but only with an explicit model of how preferences evolve in time, which falls under the rubric of model-based approaches considered more fully below.) This stationarity assumption is clearly violated in at least two situations. The first is the case of satiation. Suppose, for example, that a subject chooses repeatedly between  $n$  pieces of chocolate and  $m$  pieces of apple. The marginal value of an additional piece of chocolate may well depreciate after having chosen chocolate several times in a row. Such stimulus-specific satiety effects (still being hungry overall, but being tired of chocolate in particular) can distort indifference points, making the inference of decision variables less reliable. (Some clever researchers have actually turned stimulus-specific satiety effects to their advantage, counting on being over-fed on a particular food item to devalue it (e.g., Kringelbach *et al.*, 2003) – but more often than not satiety or, more generally, interactions between outcomes on different trials is an annoyance.) In humans, this difficulty can be alleviated by delivering a reward for one and only one of the subject's choices – this trick is commonly employed in experimental economics to (1) avoid having to pay subjects large amounts of money for performing many trials and (2) avoid the risk that subjects will attempt to assemble a portfolio or advantageous combination of options over many trials. Subjects are typically warned that they will receive no rewards during the experiment, but that one trial will be randomly selected at the end of the experiment and their decision on that trial honored. Unfortunately, this trick is unavailable to anyone working with non-human animals, and thus is of little use to neurophysiologists.

p0160 Dynamic environments present a second setting in which the assumption of stationarity implicit in

simple revealed preference becomes problematic. Increasingly, researchers have sought to enrich their experiments by varying contingencies between actions and outcomes over time, typically without warning subjects when and how the contingencies change. This technique is common in reversal-learning tasks, exploration–exploitation tasks, some foraging tasks, and any game-theoretic tasks in which an opponent reacts to the subject's actions – see, for example, Paton *et al.* (2006), Daw *et al.* (2006), Sugrue *et al.* (2004), and Barraclough *et al.* (2004), respectively. Often the experimenter is most interested in tracking changes in a decision variable during transition periods when a subject is actively adapting to the changing environment. Because they depend on combining information over many trials to find indifference points, simple revealed-preference methods can only provide precise estimates of decision variables once the subject has reached equilibrium with the environment, and not during these interesting dynamic periods.

The limitations of incentive compatible self-report p0170 and simple revealed-preference methods render them unsatisfactory for many experiments, particularly studies that involve either dynamic environments or non-human subjects. As a result, it is becoming increasingly common for researchers to construct explicit mathematical models of the decision process to estimate decision variables in experiments not amenable to the direct methods we have already discussed.

## MODEL-BASED APPROACHES

s0060

The alternative to direct methods for estimating p0180 decision variables is the model-based approach, in which a concrete mathematical model of the decision-making process is constructed (reviewed in Corrado and Doya, 2007; O'Doherty *et al.*, 2007). Some researchers view the models they construct for this purpose as literal hypotheses about the details of neural computation, but we prefer to view them as a means to an end: providing proxy variables of use in identifying decision circuits in the brain even if the details of computation differ substantially between the model and the neural implementation. This conservative stance frees us from the necessity of demonstrating that all elements of the model are plausibly implemented in the brain, and instead allows us to focus on our primary objective – identifying neural correlates of the key decision variables.

The model includes explicit decision variables that p0190 can be calculated on every trial; these quantities can then be used as *proxies* for the state of the subject's

true internal decision variables. If these proxy variables co-vary with the subject's true decision variables, we should be able to localize the neural correlates of the subject's decision variables using the traditional correlative techniques of systems neuroscience.

p0200 The general strategy for model-based investigations proceeds in several stages (Figure 30.3), each of which we will consider in a separate section below. For the remainder of this section we will focus on the first three steps in this process: selecting a general modeling framework and style, tailoring a specific model to the experimental design and scientific question, and fitting the model parameters to behavioral data. The fourth and fifth stages of the process – validating behavioral predictions and correlating model proxy variables with neural data – are common to both model-based and direct methods, and are examined in the subsequent sections.

### s0070 Styles of Modeling

p0210 There are a great many styles of models and modeling frameworks. Rather than enumerating all of the possibilities, we will outline the most important dimensions in which modeling frameworks vary and give a few examples of frameworks that have been used successfully to model behavioral choice.

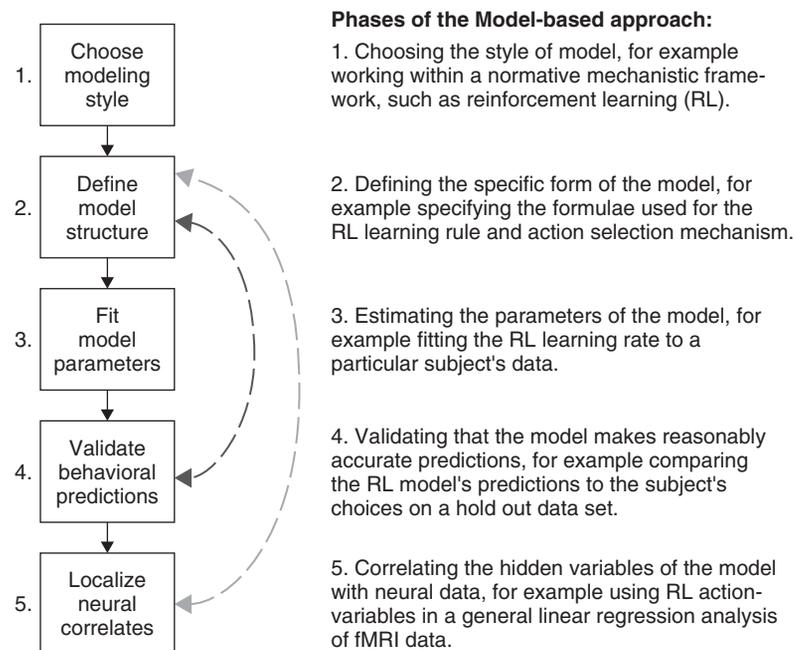
### Aggregate vs Mechanistic Models

The principle axis of variation concerns the resolution at which a model describes behavior. Some of the most influential decision models from psychology and economics deal only with average behavior – for example, Herrnstein's Matching Law (Herrnstein, 1961) or Nash Equilibria (Nash, 1950). Although these models may be very successful in describing the *aggregate* properties of choice in a wide variety of circumstances, they do not say much about individual choices. To be useful for neurobiological purposes, models need to be *mechanistic*, meaning that they should explicitly detail how internal variables are updated and used to render choices on each trial. The brain must formulate and execute each choice individually, and the value of each decision variable must therefore be specified for each trial. (Although aggregate behavior is interesting, the relevant neural mechanisms for understanding aggregate behavior are the exactly those that mediate individual choices. Aggregate behavior is after all nothing more than the cumulative *result* of the decision maker's individual choices made one at a time.) Fortunately, aggregate models can often be made mechanistic with a few simple extensions. For example, Herrnstein's Matching Law can be implemented as a local mechanistic process in which animals allocate choices based on the ratio

s0080

p0220

The Model-based approach to decision variables



f0030 **FIGURE 30.3** The five phases of a model-based approach for investigating neural decision variables. Research programs tend to progress through these stages in the presented order. Importantly, the results of both the behavioral validation and the neural correlations steps can lead to revision of the model to better capture the data (dashed lines).

of their recent reward experience (e.g. Herrnstein and Prelec, 1991; Sugrue *et al.*, 2004). As we will see later in the chapter, a key feature of mechanistic models is that they often explicitly dissociate the option evaluation stage and the action selection stage (Figure 30.5).

mechanistic-descriptive model of a decision mechanism. In doing so we will describe two additional axes of variation – the number of parameters a model requires, and the whether the model makes probabilistic or deterministic predictions.

s0090

### **Descriptive vs Normative Models**

p0230

A second important axis of variation among decision models is *descriptive vs normative*. Descriptive models seek only to *describe* the behavior as it appears in the data (e.g. Herrnstein's Matching Law), whereas normative models *prescribe* a pattern of behavior for the particular task at hand. Under the assumption that the subject is a rational agent, normative prescriptions generally aim to realize some form of optimality, such as maximizing expected reward (as in Nash Equilibria). Of course, such notions of optimality depend critically on the assumptions used to derive them, which may not align well with the constraints imposed by natural environments. For example, correctly solving for Nash Equilibria requires that a player has complete information about the precise consequences of every available action and how these outcomes depend on the actions of other agents. Even in cases where games lacking complete information are recast as Bayesian games, rich probability distributions over possible consequences and opponents' preferences are required. The correctness of the prescribed actions is sensitive to the accuracy of these large probability tables, which may themselves be intractably difficult for a decider to estimate and manipulate. Unfortunately, in many real-world situations sufficiently accurate information about what to expect from the world or our adversaries is unavailable, making the standards set by Nash unattainable for decision-making systems operating with limited information, time, and computational resources. This fragile dependence of "optimality" on specific assumptions is one of several reasons why researchers do not generally insist on limiting themselves to normative frameworks when constructing models of behavioral choice. In most cases the model that provides the more accurate account of the data is preferable, even if that model implies "rough-and-ready" decision-making on the part of the subject.

p0240

These two dimensions – aggregate/mechanistic and normative/descriptive – make up the principle axes of variation in models of decision-making. Figure 30.4 shows a table of example models in each of the four combinations of subtypes. Many other examples exist, and within each category there are additional important axes of variation. In the next two sections we will develop an example

### **Specifying the Details of the Model**

s0100

After selecting a style of model to use, we must specify its detailed mathematical form. This is largely a creative process, one that lacks a rote procedure for arriving at the best model for a particular experiment. We will consider a simple but plausible example study, and explore a few possible decision models that could be used to compute proxy decision variables and predict behavior in this hypothetical study.

p0250

In our hypothetical example experiment, we will require subjects to choose between two "lotteries." In the simplest case, a lottery is an offer of receiving reward  $x$  with probability  $p$  – for example, a 10% chance of winning \$30. (Lotteries can, of course, be much more complicated – for example, a 10% chance of winning \$30, a 5% chance of winning \$100, a 1% chance of losing \$200, and an 84% chance of winning of nothing – but here we restrict ourselves to the simplest case of winning  $x$  with probability  $p$ , and nothing with probability  $(1 - p)$ .) For our example model, we will elect to work within the mechanistic and descriptive quadrant of model frameworks (Figure 30.4, upper right quadrant). As illustrated in Figure 30.5, specifying a mechanistic model requires us to define both an option evaluation mechanism and an action selection mechanism. The mechanism selected for option evaluation is obviously critical, because this stage is directly responsible for computing the proxy decision variables that we will ultimately seek to correlate with neural activity. However, we will see below that the choice of action selection mechanism can also have substantial (if indirect) effects on the imputed decision variables that we calculate.

p0260

In the following subsection we will consider several possible option evaluation mechanisms, and in doing so explore the pros and cons of varying the number of parameters incorporated into the model. We will then go on to compare deterministic and probabilistic action selection mechanisms to marry with our preferred option evaluation mechanism in order to construct a complete model of decision-making for our lottery experiment.

p0270

### **Option Evaluation-the Pros and Cons of Parameters**

s0110

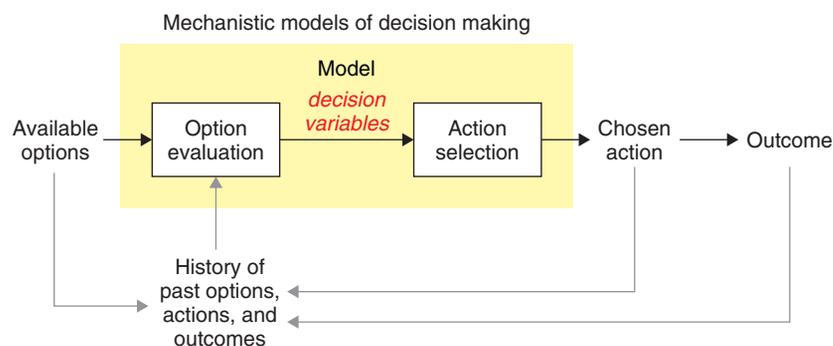
The role of an option evaluation mechanism is to convert the features of options into decision variables

p0280

Example modeling frameworks

	Aggregate	Mechanistic
Descriptive	<b>Herrstein's Matching Law</b> describes the average distribution of choices by animals foraging on multiple resource patches, without regard to whether that is the most efficient allocation of choices. (Herrstein, 1961)	<b>Linear Nonlinear Probabilistic</b> models are three-stage feed-forward models that use a linear transformation to construct decision variables and a simple nonlinearity to make probabilistic predictions of choice. (e.g. Corrado et al. 2005)
Normative	<b>Nash Equilibria</b> are stable points of competitive games, prescribed sets of strategies for which no player can improve his average payoff by unilaterally changing his behavior. (Nash, 1950)	<b>Reinforcement Learning</b> models are simple mechanisms that update the values of states or actions based on past experience, converging to optimal values over time. (Sutton & Barto, 1988)

f0040 **FIGURE 30.4** Examples of models in four corners of the descriptive/normative and aggregate/mechanistic space. Many other examples exist in each category, and the boundaries between categories may not be as crisp as implied. For example, as mentioned in the main text, Herrstein's matching law can be made more mechanistic relatively easily. Similarly, some forms of reinforcement learning (e.g., Q-learning) can be shown to be optimal in restricted environments (stationary environments that allow the agent infinite time to explore), but their optimality cannot be guaranteed under more realistic conditions.



f0050 **FIGURE 30.5** Mechanistic models of decision-making. Mechanistic models of decision-making agents often have two separable parts: an option-evaluation stage and an action-selection stage. The option-evaluation mechanism responds to immediately available options, but may also incorporate information from the past (e.g., whether an option previously delivered a positive outcome). The action-selection mechanism weighs the desirability of the available options and renders the final decision. Whether that chosen action ultimately results in a positive or negative outcome depends on factors external to the agent.

that feed the action selection mechanism. For our example experiment, this amounts to defining the subjective desirability of lotteries  $A$  and  $B$  in terms of their probabilities of payoff  $p$ , and the dollar amounts of their payoffs  $x$ :

$$(D_A, D_B) = g(p_A, x_A, p_B, x_B) \quad (30.2)$$

p0290 It seems intractable to consider all possible functions  $g$ ; we must therefore make some simplifying assumptions. Choosing which assumptions to impose is part of what makes defining a behavioral model both challenging and scientifically interesting. One common, if relatively strong, assumption is that options  $A$  and  $B$  are evaluated independently and by identical mechanisms. This greatly reduces the space of functions we are considering:

$$D_i = f(p_i, x_i) \quad i \in \{A, B\} \quad (30.3)$$

In choosing the function  $f$ , the most important axis of variation is the number of parameters we use to describe the function, and thus our model. At one end of the spectrum, *parameter-free* models assert that  $f$  (and thus the behavior of the subject) can be described by a mathematical function of our choosing, that does not need to be tuned to fit the data at hand. Relaxing this assumption a little, *parametric models* assume a very specific functional form for  $f$ , but include a small number of free parameters that can be adjusted to fit the behavioral data of individual subjects. At the far end of the spectrum, *reconstructed models* take this flexibility one step further by making relatively modest assumptions about the form of  $f$  and attempting to estimate  $f$  directly from the data – in essence, by fitting a very large number of parameters to the data. To appreciate some of the advantages and the disadvantages of these different model types, we consider some concrete possibilities in each category. (Rather

p0300

than three discrete categories, this is best thought of as a continuum, ranging from zero parameters at one extreme to a very large number of parameters (perhaps hundreds) at the other.)

p0310 A parameter-free model offers the simplest starting point. Classical economic theory tells us that the worth of a lottery is just the expected monetary value of its outcome, and so we might hypothesize that the desirability of option  $A$  is proportional to the expected value:

$$D_A \propto EV_A = p_A x_A \quad (30.4)$$

where  $p_A$  is the probability that the lottery will pay amount  $x_A$ . In the parameter-free approach, no further specification of the model is necessary. We can compute  $px$  for any option we offer, and search for neural activity that covaries with these putative values. We do not even need to determine the constant of proportionality between  $D_A$  and  $EV_A$  because its value has no effect on statistical measures of correlation. Unfortunately though, we can rarely write down parameter-free accounts of behavior that fit real choice data with any precision. Indeed, given the many factors at play in biological systems, it is rare that behavior can be reduced to an elegant mathematical formulas you might find on a T-shirt. This is not physics, after all.

p0320 In general, incorporation of even a few additional degrees of freedom improves our model's ability to describe behavior considerably. Such, *parametric models* express a strong hypothesis about the form of the decision variable, but allow individual variation in how decision variables are computed – variation that is characterized and absorbed by a small number of descriptive terms. For example, if we found that our parameter-free model (equation 30.4 above) did not fit the behavioral data very well, a neo-classical economist might suggest that desirability should be proportional not to the raw expected value but instead to the expected *utility* of an option:

$$D_A \propto EU_A = p_A U(x_A) \quad (30.5)$$

where  $U$  is the individual subject's utility function. The utility function can capture the fact that equal increments in the size of a reward may not be equally desirable to the subject – for example, some subjects may work hard to win \$1 as opposed to zero, but exert little effort to win \$101 as opposed to \$100. If we assume that a reasonable utility function should be monotonically increasing and free of inflection points, we might consider using a power law to characterize its shape:

$$U(x) \approx x^\alpha \quad (30.6)$$

where values of  $\alpha$  less than one describe concave functions and values of  $\alpha$  greater than one describe convex functions. This leads to a simple parametric model of desirability:

$$D_A \propto p_A x_A^\alpha \quad (30.7)$$

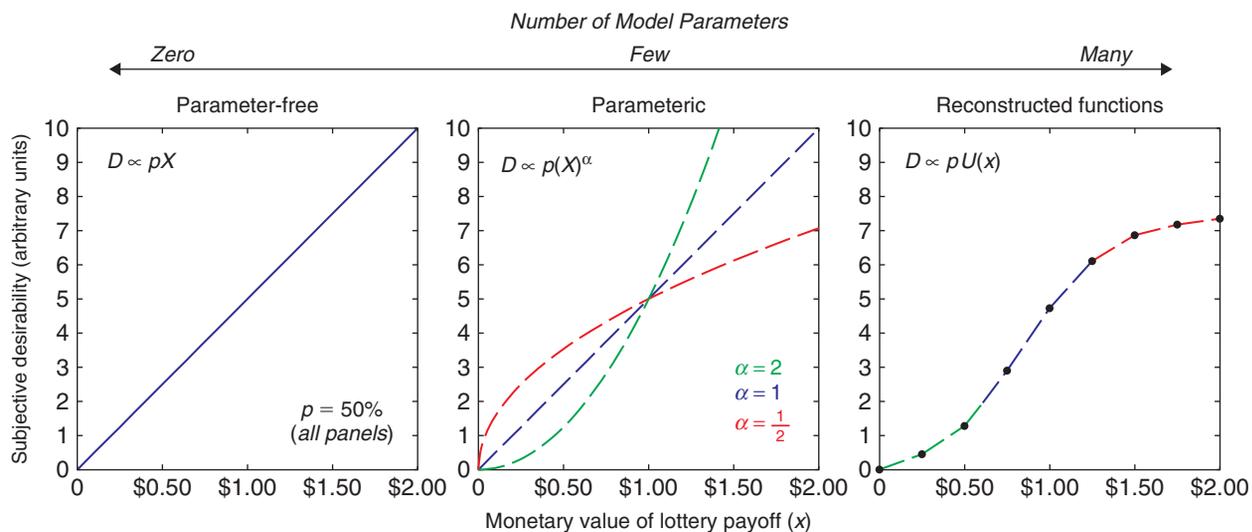
We can fit this parametric form directly to behavior; p0330 fine-tuning the value of  $\alpha$  for different subjects may allow for a more accurate description of each subject's behavior. (Our discussion of how to model lottery valuation with a utility function is simplified for didactic purposes; for a serious treatment of the problem see, for example, Holt and Laury, 2002.) We will address methods for fitting parameters to data later in the chapter.

*Reconstruction* techniques take this idea a step further. p0340 Instead of assuming a particular functional form for  $U$  and fitting the parameter  $\alpha$  to the data, we might try to measure the shape of each subject's utility function directly from the data, and then use this estimated utility function  $\hat{U}$  in our model:

$$D_A \propto p_A \hat{U}(x_A) \quad (30.8)$$

This approach is considerably more ambitious p0350 than the single parameter model described by equation (30.7). In essence, the "reconstruction" approach involves estimating a large number of parameters (for example, by expanding the function  $\hat{U}$  as a polynomial), and therefore requires much larger data sets and clever experimental designs to be tractable (see Corrado *et al.*, 2005, and Lau and Glimcher, 2005, for successful examples of the reconstruction of subjective functions, though not utility functions over lotteries, as would be needed for this example experiment). Notice that equation (30.8) does not endeavor to reconstruct the two-dimensional function  $f$  in equation (30.3) in its entirety. In general, reconstruction models still employ simplifying assumptions to reduce the problem to a manageable size. In our case, we assumed the option evaluation mechanism to be linear in its dependence on  $p$  – a relatively strong claim common to all three of the models we have considered. While this assumption is common, there is nothing sacred about it. It might be that assuming a linear dependence on  $x$  would be a better model of behavior, i.e.  $D = H(p)x$ , rather than  $D = pU(x)$ . With sufficiently large data sets we might even get away with assuming only that the dependence on  $p$  and  $x$  is separable, i.e.  $D = H(p)U(x)$ . Any of these would be far easier than reconstructing the full function  $D = f(p,x)$ .

Figure 30.6 shows a graphical representation of the p0360 option evaluation mechanisms we have considered. Moving from left to right, the number of assumptions decreases, while the number of parameters and



f0060 **FIGURE 30.6** Example option evaluation mechanisms for the subjective desirability of a lottery option with varying numbers of model parameters. The model at the left has no free parameters, and hypothesizes that the subjective desirability of a lottery is directly proportional to its expected monetary layout. The model at the center uses a single free parameter to characterize the value of a lottery as either an accelerating (green) or decelerating (red) function of possible payoff. The reconstructed function at the right uses a large number of free parameters to characterize the specific shape of the relationship between payoff and subjective desirability, allowing for richer descriptions of behavior at the cost of far greater data demands.

flexibility of the model increases. Unfortunately, with increasing flexibility comes a steadily rising demand for data to fit the increasing number of parameters. Not surprisingly, most successful research programs take the middle path – the parametric approach – avoiding both the extreme rigidity of the parameter-free models and the extreme data demands of function reconstruction.

p0370 The three example models comprised of equations (30.4), (30.7), and (30.8) provide explicit formulae for constructing decision variables that capture the desirability of a lottery option. Each example model is a reasonable candidate for the option evaluation process diagrammed on the left side of Figure 30.5, and satisfies our principal goal of explicitly specifying the putative decision variable. However, to fit the parameters of the model to data and to validate the model's behavioral predictions, we must be able to compare the subject's actual choices with choices predicted by our model. To make this comparison, we must complete our model by adding a concrete action selection mechanism, as diagrammed on the right-hand side of Figure 30.5.

#### s0120 **Action Selection: Deterministic or Probabilistic Predictions**

p0380 The role of the action selection stage is to convert putative decision variables into concrete predictions about choice. Because the action selection stage is downstream of the decision variables, it has less

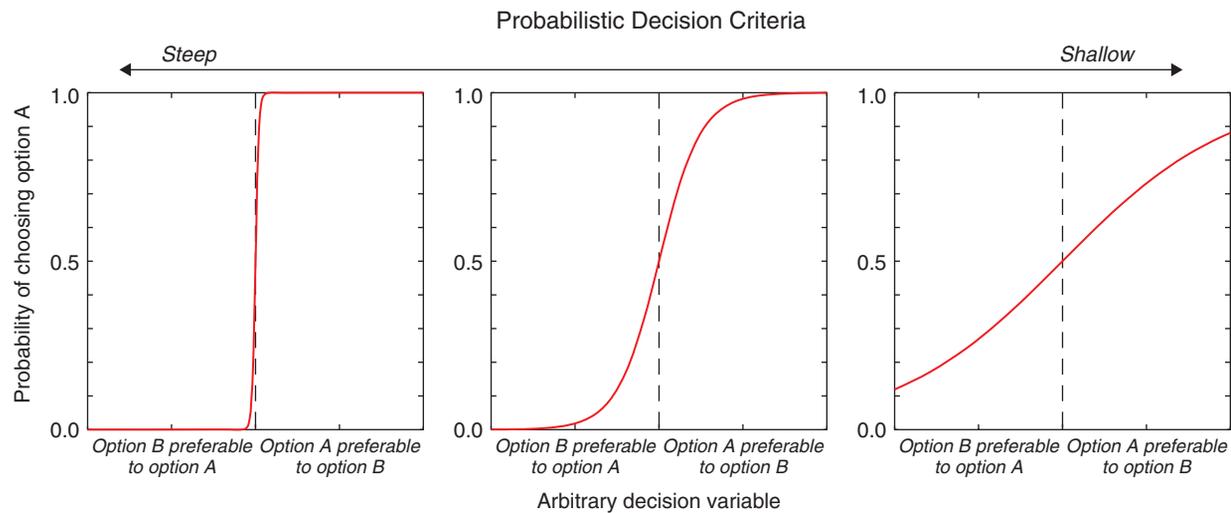
effect on our imputed decision variables than does the option evaluation stage. It exerts some influence, however, because it impacts the *parameters* of the option evaluation stage that fit the data best – thus affecting the computation of putative decision variables via changes in those parameters. Moreover, it has as critical role in our overall assessment of how well the model predicts choice, and thus our level of confidence in the hypothesized decision variables.

Action selection mechanisms come in two basic flavors: those that make *deterministic predictions* and those that make *probabilistic predictions*. By far the most common deterministic mechanism is to predict that the subject will choose the option with the greatest subjective desirability (for those unfamiliar with the *argmax* notation, this equation states that *chosen* is set to the argument *i* that maximizes the value of the expression  $D_i$ , e.g.  $chosen = A$  if  $D_A > D_B$ ):

$$chosen = \arg \max_i (D_i) \quad (30.9)$$

This action selection mechanism is extremely popular in economics, offering the advantages of simplicity and clear interpretability.

In truth, however, subjects rarely choose the “best” alternative in a deterministic manner. For this reason, many models therefore incorporate a probabilistic selection criterion: predicting simply that the *probability* of choosing an option increases with its desirability. Sigmoidal functions are generally used to implement



f0070 **FIGURE 30.7** Probabilistic decision criteria. Probabilistic decision criteria vary in how steeply they increase as a function of the desirability of an option. For example, the steep criterion at the left predicts that the probability of choosing option A is near 1 when option A is only slightly more desirable than option B. This sharp discrimination between the relative desirability of options A and B approximates the behavior of a simple deterministic function that always selects the most desirable option. The sigmoidal function at the center shows a smoother transition between exclusively choosing A or B. The gradual function at the right biases choices in favor of the more desirable option, but never chooses either option exclusively.

probabilistic selection criteria, and one of the most commonly used is the *softmax* function:

$$p_{\text{chose}}(i) = \frac{\exp\{\kappa D_i\}}{\sum_j \exp\{\kappa D_j\}} \quad (30.10)$$

where  $\kappa$  describes the steepness of the relationship between desirability and probability. Figure 30.7 shows the shape of the softmax function for different values of  $\kappa$ . The left-hand panel depicts a function that is so steep as to effectively choose the more desirable option 99.9% of the time. In contrast, the right-hand panel illustrates a function that introduces only modest biases in favor of the more desirable option, even for large differences in option desirability.

p0410 Probabilistic action selection stages allow for the possibility of noise or variability in choice behavior, and therefore encompass a far greater range of possible decision mechanisms – but, critically, they do so at relatively little cost. At the limit, probabilistic models with very steep decision criteria subsume deterministic models, accommodating behavior even when the data are best described by a nearly perfect mapping from desirability to choice.

p0420 To define a complete choice model, we can combine either of these example decision functions, equations (30.9) or (30.10), with any of our three example evaluation functions, equations (30.4), (30.7), or (30.8). For

example, using equations (30.7) and (30.10), we can write down one possible model as:

$$p_{\text{chose}}(i) = \frac{\exp\{\kappa p_i x_i^\alpha\}}{\sum_j \exp\{\kappa p_j x_j^\alpha\}} \quad (30.11)$$

This equation describes the probability that the subject will choose lottery  $i$ , which has probability  $p_i$  of paying  $x_i$ , given the parameter  $\alpha$  describing the concavity of the subject's utility function and the parameter  $\kappa$  describing the steepness of the subject's decision function.

### Parameter Estimation

s0130

Having seen that most useful models will involve some number of parameters, we now consider briefly how best to estimate those parameters from the data. In general, this is done by optimizing some *objective function*, a measure of the goodness of fit between the model and the data. Essentially, this means comparing the actual decisions made by the subject to the decisions predicted by the model, and finding the set of parameters that minimizes the discrepancy between the two.

p0430

Having specified a model that makes concrete predictions about choice – for example, equation (30.11) – we can find the model parameters that best describe

p0440

a subject's behavior by optimizing our objective function. The two most common techniques are to maximize the likelihood of the data given the model's predictions

$$L = \prod_t pchase_t(x_t) \quad (30.12)$$

or to minimize the squared error between the data and the model's predictions

$$E = \sum_t (1 - pchase_t(x_t))^2 \quad (30.13)$$

where, in both cases,  $pchase_t(x)$  is the model's predicted probability that the subject will choose option  $x$  on trial  $t$ , and  $x_t$  is the option actually chosen by the subject on trial  $t$ . Maximum likelihood (ML) and minimum squared error (MSE) parameter estimates are often very similar, and rarely is there a strong case *against* using either method. MSE parameter estimates are often easy to compute, and can accommodate both deterministic and probabilistic models. On the other hand, many view ML as the most natural sense of "goodness of fit" for probabilistic models. Moreover, in cases where there is a strong *a priori* reason to give preference to some choices of model parameters over others, ML can be extended into more advanced *Bayesian* methods (e.g. Samejima *et al.*, 2004). The details of the procedures for actually arriving at ML, MSE, or Bayesian parameter estimates depend on the details of the model being employed and are beyond the scope of this chapter, but a great many references and software tools are available (e.g. van der Heijden *et al.*, 2004).

s0140

## VALIDATING BEHAVIORAL PERFORMANCE

p0450

Whether we arrived at our putative decision variables through one of the direct methods or using the model-based approach, both outlined earlier in this chapter, it is critical to assess whether these decision variables make sound behavioral predictions. We are unlikely to achieve our ultimate goal – identifying neural correlates of hidden decision variables in the brain – if our postulated decision variables do not describe observed behavior accurately.

p0460

For those who have taken a model-based approach, assessing behavioral performance is easy: compare the choices (or probability of choice) output by the model with the subjects' actual choices and look for discrepancies. Putative decision variables arrived at through

direct methods require some additional machinery to make their predictions testable. In essence, they require an action selection mechanism just like those outlined above to transform their raw decision variables into concrete predictions about actions. Often, for the sake of simplicity, this is achieved using the deterministic "choose-the-best" rule as an implicit action selection stage (equation (30.9)), but there is no reason why a probabilistic mechanism could not be used.

## Simple Predictive Performance

s0150

p0470

The most obvious metric of a model's performance is its ability to predict a subject's choices. With a few modifications, the objective functions that we employed to fit models earlier in the chapter (see "Parameter estimation") can also be used to derive metrics of predictive performance. Clearly, the same data should not be used to define the decision variables *and* to assess the accuracy of their predictions. Asking a model to predict data that were used to set its parameters conveys an unfair advantage, and would lead to inflated confidence in the model's performance. Thus, separate subsets of the data must *always* be used to fit parameters and assess predictive performance. (The roles of the subsets may be exchanged, or even rotated in more complex schemes, and the results combined to produce so-called *cross validated* estimates of predictive performance; see Hastie *et al.*, 2001.)

p0480

For action selection mechanisms that make probabilistic predictions, the likelihood function given in equation (30.12) is by far the most commonly used metric of performance. For data sets of any significant size, the raw likelihood will be an absurdly small numerical value, because it is obtained by multiplying a long list of numbers, all less than one. For this reason, most researchers report (and indeed optimize) the log-likelihood, by taking the logarithm of both sides of the equation:

$$\text{Log}L = \sum_t \log pchase_t(x_t) \quad (30.14)$$

where, as before,  $pchase_t(x)$  is the model's predicted probability that the subject will choose option  $x$  on trial  $t$ , and  $x_t$  is the option actually chosen by the subject on trial  $t$ . For the model-based approach, when comparing the performance of models having different numbers of parameters, heuristics like Bayesian Information Criterion (BIC) or Akaike's Information Criterion (AIC) can be used to augment equation (30.14) with a term that penalizes models for having additional parameters (see Burnham and Anderson,

2002, for a tutorial). (These metrics include a penalty for additional parameters because we would expect any model with more parameters to fit better. Unfortunately, BIC and AIC are only heuristics for making this correction.)

p0490 Likelihood metrics like equation (30.14) cannot be applied to action selection mechanisms that make deterministic predictions. Fortunately, error metrics like equation (30.13) are available as ready substitutes. For deterministic predictions, equation (30.13) reduces to the percentage of the subject's actual choices that are correctly predicted by the model, a simple but very intuitive metric (e.g. "our model correctly predicted 98% percent of subjects' decisions"). While commonly used for deterministic predictions, this metric is rarely used for probabilistic models because it takes no account of the confidence with which predictions are made.

### s0160 **Establishing Upper and Lower Bounds on Predictive Performance**

p0500 The metrics of predictive accuracy described above can quantify how well a particular model describes a specific data set, and thus can be used to compare the performance of competing models. (The same metrics can be used to evaluate decision variables arrived at through direct (rather than model-based) approaches; however, as mentioned earlier, the directly inferred decision variable must be coupled with a particular action selection mechanism before its predictive accuracy can be assessed.) However, thus far we lack a "gold standard" that specifies in principle how well a model could describe the data. More generally, we would like to establish bounds on both the low and the high ends of performance, so as to assess any specific model relative to random and ideal performance.

p0510 We can establish a reasonable lower bound on performance by applying a metric of predictive performance to a set of "random" predictions. If a model cannot out-perform a random noise generator, clearly we should discard it and search for something more promising. (There is some liberty in how we define these "random" predictions. The simplest option is to make *choices* that are random: a dummy set of predictions,  $p_{dummy}$ , uniform and independent of the options offered:  $p_{dummy}_t(k) = 1/m$  for all trials  $t$  and options  $k$  when choosing between  $m$  alternatives. Another possibility is to make *random predictions*, i.e. drawn at random from a uniform distribution:  $p_{dummy}_t(k) = r$  and  $[0,1]$ . For the purpose of comparison with a single particular model, there is also the possibility of using *randomly shuffled predictions*, that is to use predictions that are drawn at random from

the predictions made by the model under scrutiny:  $p_{dummy}_t(k) = p_{chosed_j(k)}$ , so that the  $t$ -th dummy prediction is set to be the comparison model's  $j$ -th prediction for a randomly chosen trial  $j$ .) Establishing a reasonable upper bound on predictive performance is more difficult. Predicting every choice perfectly is, of course, an absolute upper bound, but an unreasonable expectation for any sizeable data set. More sensibly, for the data under consideration, we would like to be able to compare the predictive performance of the model being considered with the maximum achievable performance of any compact model.

An elegant approach to calculating maximum achievable performance was suggested to us by p0520 William Bialek of Princeton University. For any model  $M$  using the inputs  $X$  to predict choices  $C$ , we can establish an upper bound on the quality of the predictions using Shannon's mutual information:

$$I(C; M(X)) \leq I(C; X) \quad (30.15)$$

Put into words, the information about subject choices contained in the model's predictions cannot exceed the information about those choices contained in the model's inputs (Figure 30.8). Models cannot add Shannon information; they can only distill it. This realization offers a practical method for calculating an upper bound on predictive performance. Given particular inputs (for example, descriptions of the options A and B in terms of probability and payoff) we can quantify how well *any* model relying on these inputs could possibly do by computing  $I(C; X)$ . If a subject behaves consistently, making similar patterns of choices for repeated presentations of identical option sets, the upper bound will be high, perhaps approaching (if never reaching) the elusive goal of perfect

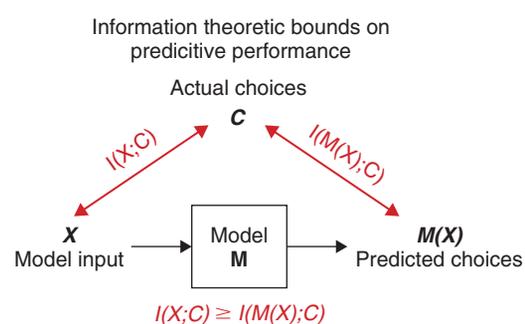


FIGURE 30.8 Information theoretic bounds on predictive performance. Information theory provides a rigorous bound for the predictive performance of behavioral models. Consider all models  $M$  that use the variables  $X$  as inputs. The mutual information between the model's predictions,  $M(X)$ , and the actual choice data,  $C$ , cannot exceed the mutual information between  $X$  and  $C$ . Thus,  $I(X; C)$  provides an upper bound on performance for all models that use  $X$  to predict  $C$ .

prediction. If, on the other hand, the subject behaves erratically, making different choices for repeated presentations of identical option sets, the upper bound on model performance will be lower. For example, in our own dynamic foraging experiments, the history of a subject's recent rewards and choices contained nearly 0.50 bits of predictive information about their next choice, but varied somewhat from subject to subject. This set an upper bound on the performance of any model that used reward and choice history as an input to predict a subject's behavior. The mechanistic models considered in Corrado (2007) yielded decision variables that retained between 50% (the worst models tested) and 90% (the best models tested) of this available information.

p0530 This procedure for establishing upper bounds on performance has a serious limitation: the size of the data set required to estimate  $I(C; X)$  is typically much larger than can be gathered from individual human subjects in a reasonable amount of time. The reason is that the probability table required to map out the function  $I(C; X)$  over all possible inputs  $X$  and range of choices  $C$  is usually very large. Our data set, for example, included tens of thousands of trials per subject, and still did not permit us to specify  $I(C; X)$  completely. Thus, though the idea shows great promise, the problem of limited data sets currently diminishes its practical generality.

### s0170 Generative Performance

p0540 In contexts where successive decisions are made independently of one another (i.e. the subject would have made the same decision on a particular trial even if the trials had been presented in some other order), evaluating average predictive performance is frequently sufficient to quantify behavioral performance. However, in contexts where the environment (or the state of the subject) has interesting temporal dynamics, an additional assessment is needed: the ability to *generate* realistic patterns of behavior in a simulated experiment. (Dynamic environments include any task where contingencies between actions and outcomes change over time and must be inferred by the subject, as is common in foraging tasks, reversal-learning tasks, exploration–exploitation tasks, and nearly all game-theoretic tasks. Dynamics internal to the subject refer to any setting in which the subject's next choice depends either on previous actions or previous outcomes, as in contexts involving reinforcement learning or “reactive” player models.)

p0550 To highlight the distinction between *predictive* and *generative* performance, consider a television weatherman who everyday forecasts, “today's weather will be

the same as yesterday's.” Because of temporal correlations in weather patterns, this algorithm actually *predicts* tomorrow's weather correctly on an impressively high proportion of days. However, because the algorithm incorporates no understanding of the mechanisms underlying weather production, it will fail miserably at *generating* realistic patterns of weather change over time. Thus, adequate predictive performance (as described in the previous section) is no assurance of generative success.

Metrics for evaluating generative performance are more difficult to define than their counterparts for assessing predictive performance. The principle, however, is simple: compare the dynamics in the real data with those of synthetic data generated using the model, and look for discrepancies. In practice, of course, we should focus on dynamics pertinent to the particular task. For example, in a reversal-learning task with an abrupt un signaled change in reward contingencies, we might look at the time-course of the subject's adaptation to the new conditions (e.g., Kennerley *et al.*, 2006). On the other hand, in a foraging task we might compare the distribution of the lengths of time spent at each foraging patch (Corrado *et al.*, 2005). Regardless of the specifics, the ideal model will produce synthetic data that are statistically indistinguishable from real data. We are not aware of any simple formula that specifies the precise meaning of “statistically indistinguishable” for the completely general case. Instead, evaluating generative performance is more akin to a “Turing test” for behavioral models: if experimenters using any of the behavioral measures of interest cannot tell which data are generated by the model and which are generated by the subject, then the model is presumably a good one. (Alan Turing made a famous proposal that a reasonable test for artificial intelligence would be to stage what amounts to an instant messaging text conversation, and if humans were unable to distinguish between talking to another human or to a machine, then that machine could be said to be intelligent (Turing, 1950).)

p0560

### THE QUEST FOR NEURAL CORRELATES

s0180

The traditional program of experimental psychology and economics ends here, with the extraction of internal decision variables that have demonstrable predictive power. Neuroscience and neuroeconomics aim to go further: to use these putative decision variables to elucidate decision-making circuitry within the brain. The core idea is to use these inferred variables as proxies for the subject's true decision variables in

p0570

a correlative study (Figure 30.2b). We have already alluded to studies in which decision variables recovered through the direct methods described above have been used to probe underlying neural signals (see, for example, Plassman *et al.*, 2007, for incentive compatible direct report in human fMRI; Padoa-Schioppa and Assad 2006 for revealed preference in monkey electrophysiology). In this section we consider three examples of the integrated model-based approach and its successful application to the analysis of neural signals (Figure 30.3).

p0580 The three example studies differ in the physiological signals measured, the species studied, and the specific form of the model used to capture the behavioral data. However, all employ tasks in which option values are uncertain and change unpredictably through time. In each study the authors built a mechanistic model of task behavior that had distinct option evaluation and action selection stages (Figure 30.5), and used the model's internal decision variables to identify neural correlates. Finally, and perhaps most importantly, in each case the authors ultimately refined their initial model(s) in response to subsequent behavioral or neural findings. This last point emphasizes the inherently iterative nature of this process (dashed lines in Figure 30.3), and the fact that a behavioral model, regardless of its normative elegance or predictive power, should not be regarded as anything more or less than a working hypothesis (but see Chapter 3 in this volume for an alternative view).

### s0190 Matching Behavior and Value-based Action Selection in Area LIP

p0590 In one of the earlier examples of the model-based approach (Figure 30.3), the authors of this chapter employed a Linear-Nonlinear-Poisson (LNP) framework to build a parametric model of the behavior of rhesus monkeys engaged in a dynamic foraging task (Sugrue *et al.*, 2004). This model was inspired by, and amounts to, a mechanistic implementation of Herrnstein's Matching Law (Herrnstein, 1961). At its evaluation stage the model uses simple exponential filters to compute a weighted average of recent rewards from each option, while at its action selection stage it replicates the matching law by using divisive normalization to transform these value estimates into a probabilistic output appropriate for selecting an operant response.

p0600 Despite its simplicity, this model was successful both in predicting the animals' choices and in generating realistic simulated behavior. Furthermore, the model's trial-to-trial estimates of value revealed a

monotonic encoding of option value in the firing rates of individual neurons in cortical area LIP, a region known to be involved in the high-level planning of eye movements. Interestingly, while option color was the cue to value in this task, the map of option value in LIP was in spatial coordinates, appropriate for the probabilistic selection of an appropriate action on the current trial. This transformation from color-based to spatial encoding of value suggested to the authors that activity in LIP reflects, and might even effect, the action selection mechanism that maps value to action in this task (Figure 30.5).

Corrado *et al.* (2005) later took advantage of the large datasets collected in this experiment to move from a parametric model to full function reconstruction (Figure 30.6) of the option evaluation and action selection stages directly from the data. The resulting model differed in a number of respects from the original parametric model, the evaluation stage being better characterized by a double rather than a single exponential, and the action selection stage by a differential (softmax) rather than a divisive operation. Using the validation techniques discussed above, the authors showed that the reconstructed model provided a better account of both animal behavior and LIP firing rates than the original parametric model, in spite of the fact that it no longer bore any formal resemblance to the original matching law. This result highlights the potential for the choice of modeling style – in this case, parametric versus reconstructed – to influence the conclusions of the entire behavior-model-physiology process.

### fMRI Correlates of Reinforcement Learning Models

The next two studies used Reinforcement Learning (RL) as the basis for the evaluative stage of their behavioral models (Sutton and Barto, 1988). Accordingly, in addition to variables that represented the learned values of competing options, these models included a RL prediction-error term that quantified the difference between the value of the chosen option and the outcome of that choice. The correspondence between the error signals postulated by RL models and the phasic responses of midbrain dopamine neurons recorded in simple conditioning tasks has led to the proposal that RL might reflect the literal mechanism through which the brain learns the values of states and actions (see, for example, Schultz *et al.*, 1997; also Part 3 of this volume). In that regard, the following studies are of particular interest because they are among the first neuroimaging studies to

examine RL in choice situations where subjects are actively engaged in learning option values and using those values to guide behavior.

p0630 Daw, O'Doherty and colleagues (2006) employed a well-characterized behavioral paradigm called the "N-armed bandit" task. Their subjects made repeated choices among four slot machines, each paying out with an undisclosed probability that varied unpredictably through time. This design placed their human subjects in a situation similar to that of the monkeys in the matching task described above: challenged with estimating the machines' relative values, and using those estimates to decide which machine to choose on each trial. An RL algorithm provided the value estimates for the authors' behavioral model, while a softmax operation comprised the action selection stage (softmax having been selected after considering a variety of different rules using maximum likelihood techniques).

p0640 The authors used the internal variables of the resulting model as regressors to analyze fMRI BOLD measurements made while their subjects performed the task. The task-related fMRI signals identified in their analysis initially seemed to confirm what other groups had reported previously: BOLD signals in ventromedial prefrontal cortex (vmPFC) correlated with the model's estimates of the relative value of the chosen option, while signals in the striatum correlated with the error term of the RL algorithm (Knutson *et al.*, 2000; McClure *et al.*, 2003; O'Doherty *et al.*, 2003; see also Chapter 24 in this volume).

p0650 If it had ended there, this result would have marked an important demonstration of RL signals within the brain in the context of a task in which those signals have true behavioral relevance. However, the authors extended their analysis in a further and highly revealing direction. Using the predictions from their model's action selection stage, they classified each choice made by their subjects as either exploitative (choosing the machine assigned the highest value by the model) or exploratory (choosing a lesser valued option). They then searched for regions where the BOLD signal differentiated between these two classes of trials, identifying areas of anterior frontal and posterior parietal cortex that were preferentially active during exploratory decisions. This result suggested the involvement of distinct brain regions that help guide exploratory actions, above and beyond the signals expected in the most naïve forms of RL.

p0660 Recent work by Hampton and colleagues (2006) extended this result in an experiment that is among the first in the field to use neural signals to explicitly test competing behavioral models. These investigators studied behavior and fMRI BOLD responses in the context of a simple probabilistic reversal-learning

task. Human subjects could improve their performance through their knowledge of the structure of the task, which included quasi-periodic reversals in the identity of the more valuable of two available options. Appropriately, the authors found that a state-based Bayesian Markov model that incorporated information about this higher-order task structure, provided a better account of subjects' behavior than simple RL models that were blind to such contextual information. Importantly, the decision variables at the evaluation stage of these two competing models sometimes diverged widely in their predictions, allowing the authors to show that BOLD signals in the striatum and vmPFC – two areas associated with RL signals – more closely tracked the update and value signals derived from the abstract states of their Bayesian model. To our knowledge, this is the first neuroeconomics study to use neural signals to distinguish between two candidate models of decision-making. However, as our understanding of neural mechanisms matures and our models increase in sophistication, such hypothesis-driven approaches should become routine.

## CONCLUSION

s0210

We began this chapter with the challenge of studying the neurobiology of decision-making in settings where we lack easy access to the decision variables that guide choice. We subsequently saw how mechanistic behavioral models can offer a solution to this problem by providing reasonable proxy variables with which to correlate physiological signals. The putative decision signals identified through this approach must ultimately become the raw materials for a neurophysiological program aimed at transforming our behavioral models from preliminary sketches to plausible blueprints for the brain's actual decision-making machinery.

p0670

Such a program entails two phases. The first demands the detailed characterization of neural signals from the perspective of different hypotheses and models. As discussed previously, an active dialogue between modeling and physiology is essential for this process to converge on a representation that is true to the underlying neural signals. The second phase hinges on the critical distinction between correlation and causation. Demonstrating neural correlates of a decision variable is, in principle, straightforward; it is substantially more challenging to prove that the correlated neural activity plays a causal role in the brain's decision-making process in the manner suggested by that proposed decision variable.

p0680

p0690 Establishing causal links between neural responses and perceptual or cognitive phenomena is a fundamental challenge faced by researchers not only in neuroeconomics, but also in all of cognitive neuroscience. Historically, support for links between anatomy and function has come from patients or experimental animals with lesions restricted to the anatomic area of interest. Indeed, lesion studies first implicated vmPFC in value-based decision-making by demonstrating that damage to this region impaired performance on reward-cued reversal learning tasks (reviewed in Rolls, 2000) and other tasks in which the best choice on each trial had to be inferred from the outcomes of earlier choices (see, for example, Bechara *et al.*, 2000).

p0700 Although they have proven useful for certain purposes, lesion techniques are rather blunt tools for dissecting neural mechanism. What we want are techniques that are temporally, spatially, and functionally precise, allowing us to manipulate, selectively and reversibly, the activity of a restricted population of neurons while leaving the rest of the system intact. Toward this end, transcranial magnetic stimulation (TMS) has been used in humans to induce transient inactivations of particular brain areas, permitting inference about their involvement in a given perceptual or cognitive function (reviewed in Hallett, 2007). In animals, similar reversible inactivations can be achieved through the local application of drugs (such as muscimol or lidocaine) which briefly silence neural activity in a restricted region (e.g. Hikosaka and Wurtz, 1985; McPeck and Keller, 2004).

p0710 Causal relationships between neural activity and cognitive function can also be tested using intracortical microstimulation (ICMS) – a technique that has relatively high spatial, temporal, and functional resolution. The idea behind this technique is simple: characterize the response properties of a local population of neurons, place an animal in a situation in which the signals from that population have direct relevance to behavior, and then ask whether microstimulation of the neurons influences behavior in a predictable manner. Pioneered in the field of perceptual decision-making (e.g. Salzman *et al.*, 1990), this technique has more recently been applied successfully to the study of visual attention, face discrimination, target selection, and reward-based learning (for a review, see Cohen and Newsome, 2004). The one caveat to this approach is that its functional specificity relies on a certain level of organization in the underlying neural architecture. Specifically, the technique has proved most successful when applied in brain areas or structures in which neurons with similar response properties are spatially clustered, presumably allowing for microstimulation to inject a coherent signal into the underlying neural circuitry.

Ultimately, more powerful circuit manipulation techniques may arise from promising developments in optogenetics (for a review, see Zhang *et al.*, 2007). Using genetic manipulations, small light-activated channel proteins are inserted into a circuit of choice. The channels can be either excitatory or inhibitory, and they can be activated and inactivated with excellent temporal precision simply by shining light on the neurons containing the channels, either via direct illumination for surface structures or via fiber-optic probes for deep structures. These techniques are potentially extremely powerful, and are being developed rapidly in species that are particularly amenable to genetic manipulation (such as mice and flies). Introduction of the channel genes into non-human primates via viral vectors appears feasible, and several labs are initiating work in this area. Application to humans, though a seemingly distant prospect at the moment, might arrive sooner than we think. This fact emphasizes how important it is that human and animal studies proceed in parallel. Work in non-human primates is particularly important in this regard, as these animals can perform many of the same complex decision-making tasks as humans. Fortunately, this parallel development has been a feature of this field since its inception, and initial signs suggest that a good deal of the basic neural machinery of decision-making and valuation may be shared between species.

In neuroeconomics, we have yet to test causal links between particular neural circuits and behavioral choice. It is exciting to imagine a day in the near future when microstimulation, optogenetics, or other techniques might be used to test the hypotheses that are beginning to emerge about the functional roles of various structures. With so many of the fundamental discoveries in this field still on the horizon, and the tools with which to explore them finally becoming available, we anticipate that studying “choice” will continue to be worth the “trouble” for a considerable time to come.

## References

- Barraclough, D.J., Conroy, M.L., and Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci.* 7, 404–410.
- Bechara, A., Damasio, H., and Damasio, A.R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex* 10, 295–307.
- Burnham, K.P. and Anderson, D.R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. New York, NY: Springer-Verlag.
- Cohen, M.R. and Newsome, W.T. (2004). What electrical microstimulation has revealed about the neural basis of cognition. *Curr. Opin. Neurobiol.* 14, 169–177.

- Corrado, G. S. (2007). "Leveraging behavioral models to reveal the neural representation of value." *Doctoral Dissertation*, Stanford University.
- Corrado, G. and Doya, K. (2007). Understanding neural coding through the model-based analysis of decision making. *J. Neurosci.* 27, 8178–8180.
- Corrado, G.S., Sugrue, L.P., Seung, H.S., and Newsome, W.T. (2005). Linear-nonlinear-Poisson models of primate choice dynamics. *J. Exp. Anal. Behav.* 84, 581–617.
- Daw, N.D., O'Doherty, J.P., Dayan, P. *et al.* (2006). Cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879.
- Gold, J.I. and Shadlen, M.N. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–574.
- Hallett, M. (2007). Transcranial magnetic stimulation: A primer. *Neuron* 55, 187–199.
- Hampton, A.N., Bossaerts, P., and O'Doherty, J.P. (2006). The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J. Neurosci.* 26, 7836–8360.
- Hastie, T., Tibshirani, R., and Friedman, J. (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY: Springer.
- Herrnstein, R.J. (1961). Relative and absolute strength of responses as a function of frequency of reinforcement. *J. Exp. Anal. Behav.* 4, 267–272.
- Herrnstein, R.J. and Prelec, D. (1991). Melioration: a theory of distributed choice. *J. Econ. Persp.* 5, 137–156.
- Hikosaka, O. and Wurtz, R.H. (1985). Modification of saccadic eye movements by gaba-related substances. I. Effect of muscimol and bicuculline in monkey superior colliculus. *J. Neurophysiol.* 53, 266–291.
- Holt, C.A. and Laury, S.K. (2002). Risk aversion and incentive effects. *Am. Econ. Rev.* 92, 1644–1655.
- Kennerley, S.W., Walton, M.E., Behrens, T.E. *et al.* (2006). Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* 9, 940–947.
- Knutson, B., Westdorp, A., Kaiser, E., and Hommer, D. (2000). fMRI visualization of brain activity during a monetary incentive delay task. *NeuroImage* 12, 20–27.
- Kringelbach, M.L., O'Doherty, J., Rolls, E.T., and Andrews, C. (2003). Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. *Cerebral Cortex* 13, 1064–1071.
- Lau, B. and Glimcher, P.W. (2005). Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* 84, 555–579.
- 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.
- McPeck, R.M. and Keller, E.L. (2004). Deficits in saccade target selection after inactivation of superior colliculus. *Nat. Neurosci.* 7, 757–763.
- Nash, J. (1950). equilibrium points in n-person games. *PNAS* 36, 48–49.
- Ochsner, K.N., Knierim, K., Ludlow, D.H. *et al.* (2004). Reflecting upon feelings: an fMRI study of neural systems supporting the attribution of emotion to self and other. *J. Cogn. Neurosci.* 16, 1746–1772.
- 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.
- O'Doherty, J.P., Hampton, A., and Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. *Ann. NY. Acad. Sci.* 1104, 35–53.
- Padoa-Schioppa, C. and Assad, J.A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226.
- Paton, J.J., Belova, M.A., Morrison, S.E., and Salzman, C.D. (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* 439, 865–870.
- Plassmann, H., O'Doherty, J., and Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* 27, 9984–9988.
- Reale, R.A. and Imig, T.J. (1980). Tonotopic organization in auditory cortex of the cat. *J. Comp. Neurol.* 192, 265–291.
- Rolls, E.T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex* 10, 284–294.
- Romo, R. and Salinas, E. (2003). Flutter discrimination: neural codes, perception, memory and decision making. *Nat. Rev. Neurosci.* 4, 203–218.
- Ross, L. and Nisbett, R.E. (1991). *The Person and the Situation*. Philadelphia, PA: Temple University Press.
- Salzman, C.D., Britten, K.H., and Newsome, W.T. (1990). Cortical microstimulation influences perceptual judgements of motion direction. *Nature* 346, 174–177.
- Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science* 310, 1337–1340.
- Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.
- Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.
- Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2005). Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat. Rev. Neurosci.* 6, 363–375.
- Sutton, R.S. and Barto, A.G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.
- Turing, A. (1950). Computing machinery and intelligence. *Mind* 236, 433–460.
- van der Heijden, F., Duin, R.P.W., de Ridder, D., and Tax, D.M.J. (2004). *Classification, Parameter Estimation and State Estimation: An Engineering Approach Using MATLAB*. Chichester: Wiley.
- Zhang, F., Aravanis, A.M., Adamantidis, A. *et al.* (2007). Circuitbreakers: optical technologies for probing neural signals and systems. *Nat. Rev. Neurosci.* 8, 577–581.