

Understanding Neural Coding through the Model-Based Analysis of Decision Making

Greg Corrado¹ and Kenji Doya²

¹Stanford University, Stanford, California 94305, and ²Okinawa Institute of Science and Technology, Okinawa 904-2234, Japan

The study of decision making poses new methodological challenges for systems neuroscience. Whereas our traditional approach linked neural activity to external variables that the experimenter directly observed and manipulated, many of the key elements that contribute to decisions are internal to the decider. Variables such as subjective value or subjective probability may be influenced by experimental conditions and manipulations but can neither be directly measured nor precisely controlled. Pioneering work on the neural basis of decision circumvented this difficulty by studying behavior in static conditions, in which knowledge of the average state of these quantities was sufficient. More recently, a new wave of studies has confronted the conundrum of internal decision variables more directly by leveraging quantitative behavioral models. When these behavioral models are successful in predicting a subject's choice, the model's internal variables may serve as proxies for the unobservable decision variables that actually drive behavior. This new methodology has allowed researchers to localize neural subsystems that encode hidden decision variables related to free choice and to study these variables under dynamic conditions.

Key words: behavior; cognitive; decision; discrimination; fMRI; learning; reinforcement; reward

Classic correlative approach

Over the past 50 years, the dominant approach in systems neuroscience has been correlative: manipulate a variable of interest, compare neural responses across conditions, and attribute observed differences to changes in functional processing. This paradigm enjoys extremely broad applicability, forming the core of analyses ranging from the plotting of tuning curves in single-unit electrophysiology to the use of statistical parametric maps in functional magnetic resonance imaging. In the simplest case, when studying early sensory processing, neural activity can be correlated directly with the sensory stimuli presented by the experimenter. For example, Hubel and Wiesel (1959) observed that the firing of neurons in primary visual cortex (V1) covaried with the orientation of a bar stimulus, and so discovered columnar orientation tuning V1.

In higher-level sensory processing, however, neural activity might depend not only on the physical stimuli presented, but on the subject's perceptual experience of those stimuli. Thus, researchers realized the need to sort neural data both by stimulus condition and by a subject's behavioral response to the stimulus. For example, Newsome and colleagues trained animals to report the direction of motion of a noisy stimulus with an eye movement and then showed that the firing rates of the neurons in area MT were correlated with the reported perception of motion, even in cases in which the visual stimulus itself was ambiguous and identical (Newsome et al., 1989; Britten et al., 1996). This classic correlative paradigm has even proven useful in the study of simple

decision variables related to sensory discriminations. For example, in a similar motion discrimination task, Shadlen and colleagues have shown that the evolving activity of neurons in the lateral intraparietal area (LIP) reflects the accumulated sensory evidence in favor of one or the other alternative (Shadlen and Newsome, 1996; Huk and Shadlen, 2005).

Search for neural correlates of subjective variables

How can we extend this correlative paradigm to search for neural correlates of decision making, beyond the simple case of sensory discrimination? In general, decisions are guided by internal subjective variables whose state cannot be directly measured or controlled by the experimenter. Instead, these variables reflect the history of a subject's actions, reward experience, and perhaps even innate individual differences. And although an experimenter may hope to shape these variables over time by pairing stimuli or actions with outcomes (i.e., classical and operant conditioning), the impact that these manipulations have on a trial-by-trial basis is not immediately obvious.

Under some conditions, these internal variables are posited to be relatively constant. For example, when asked to choose between familiar food or liquid rewards, an animal's expressed preferences can be used to infer the relative value of the options. Using this approach, several groups have shown that neural activity in orbitofrontal cortex is correlated with the subjective value of offered or chosen goods (Tremblay and Schultz, 1999; Padoa-Schioppa and Assad, 2006). But how can we establish neural correlates under the more common (and potentially more interesting) scenario that a decision variable is changing over time?

The most obvious solution is to construct an explicit model of the decision process within the context of a given experimental

Received April 9, 2007; revised June 11, 2007; accepted June 12, 2007.

Correspondence should be addressed to Greg Corrado at the above address. E-mail: gcorrado@stanford.edu.

DOI:10.1523/JNEUROSCI.1590-07.2007

Copyright © 2007 Society for Neuroscience 0270-6474/07/278178-03\$15.00/0

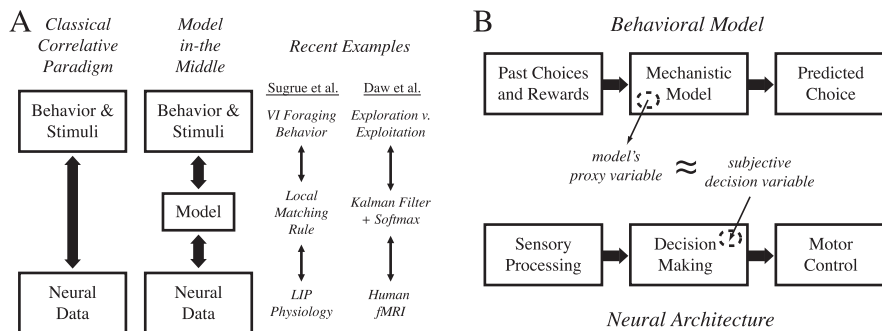


Figure 1. *A*, Left, The classical paradigm directly correlates neural data with stimuli or behavior. Center, The “model-in-the-middle” approach instead uses a mechanistic model as an intermediary, on the one hand constraining the model to describe behavior and on the other correlating internal variables of the model with neural data. Right, Two examples of triplets of behavior, model, and neural data, one from Sugrue et al. (2004) and another from Daw et al. (2006). *B*, The basic scheme for using models as intermediaries. Top, Mechanistic models of choice generally take as input the history of past choices and rewards and render a prediction of future choice as an output. Center, Internal variables computed by the model can be extracted and used as proxies for subjective decision variables hidden from the experimenter. Bottom, These proxy variables can be used to identify specific neural representations of the decision variable of interest within the brain areas supporting decision making.

task. In this way, the impact of each successive action and reinforcer on the internal decision variables can be estimated via the behavioral model. If the model is able to capture the dynamics of the subject’s actual decisions, the internal variables of this model may be suitable to serve as proxies in the search for correlates of subjective variables in neural data. Thus, breaking with the classical paradigm in which neural data are correlated directly with stimuli and behavior, in this new approach neural data are instead correlated with the internal variables of an intermediary behavioral model (Fig. 1*A*).

The spectrum of decision-making models

Some of the most influential decision models from psychology and economics deal only with average behavior [e.g., Herrnstein’s Matching Law (Herrnstein, 1961) or Nash Equilibria (Nash, 1950)]. These models describe the aggregate properties of choice in a wide variety of circumstances but do not specify how these choice patterns are realized. To be useful for our purposes, however, models must be mechanistic, meaning that they explicitly detail how internal variables are updated and used to render choices on each trial. Fortunately, aggregate models can often be made mechanistic with a few simple extensions [e.g., the “Local” Matching Law (Sugrue et al., 2004)].

Another important axis of variation among decision models is descriptive versus 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 that realizes some form of optimality, such as the maximization of cumulative rewards (e.g., Nash Equilibria). Of course, the optimality of these prescriptions depends critically on the assumptions used to derive them, which may not align well with the constraints imposed by natural environments. This is one of several reasons that the use of normative models is not stressed in this context.

Perhaps the most popular type of model used as intermediaries are those derived from Reinforcement Learning Theory (Sutton and Barto, 1998; Doya, 2007). These models are both normative and mechanistic and have great flexibility in their applicability. Rather than assuming any particular prior knowledge, these models detail a method for how an initially ignorant subject should update simple internal variables to ultimately achieve optimal performance.

Regardless of their discipline of origin, most suitable mechanistic models have free parameters that must be chosen (or fit) to capture observed behavioral data. The standard approach to model fitting in this context is maximum likelihood estimation (e.g., Lau and Glimcher, 2005). In this approach, we assume a set of parameters, run the model using the real subject’s experiences as input, and compare the actions predicted by the model with those of the subject. By systematically scanning the parameter space, we can find a set of parameters and associated time course of decision variables that best explains the subject’s behavior. It is also possible to apply more sophisticated estimation methods, such as Bayesian inference (e.g., Samejima et al., 2004), which allow estimation of time-varying parameters under certain conditions. Often there are several models

that fit the behavior reasonably well, and in these cases the ultimate performance of a model in extracting neural correlates is the best metric of its utility.

Neural recording experiments

Samejima et al. (2005) provide an example of how a computational model can be used to construct proxy variables for correlation with neural responses. The crux of the methodology is to compute an internal variable derived from an intermediary behavioral model and then leverage that variable to search for decision-related correlates in the nervous system (Fig. 1*B*). The decision variable of interest to Samejima et al. (2005) was action value, the subjective expectation of reward for a particular action. Action value is a prime example of a variable whose exact state is not immediately obvious from observable experimental parameters but can be inferred using a mechanistic behavioral model. Samejima et al. (2005) computed action values on each trial by applying a standard reinforcement learning model, Q-learning, to the history of actions and outcomes that the animal had experienced before that trial. They found that firing rates of many striatal neurons were well correlated with the computed action values, suggesting a role for this area in representing the values of competing alternatives during decision making.

The true usefulness of this paradigm, however, is most apparent when the model uncovers effects that would otherwise have gone unnoticed. In the previous study, the blockwise design of the task was such that the central result could have been described even without the use of a model. The clearest example of a case in which an intermediary model was essential to understanding neural signals comes from recent attempts to identify value-related signals in area LIP. Platt and Glimcher (1999) showed that LIP firing rates seemed to vary with reward expectation in an instructed saccade task. However, in a free-choice task, almost no variation of LIP firing rates can be seen with changing reward contingencies when data are combined across trials in the traditional way (Dorris and Glimcher, 2004; Sugrue et al., 2005). In contrast, combining these same data based on proxy variables computed by value-based decision models reveals strong value-related modulations in LIP (Sugrue et al., 2004) (see also Dorris and Glimcher, 2004). Although the magnitude of these effects rivals well known modulations of LIP activity by saccade planning and spatial attention [for review, see Andersen (1995) and

Colby and Goldberg (1999)], these value-related signals would be nearly undetectable without the use of intermediary behavioral models to direct the search.

Neuroimaging experiments

The utility of this approach extends to studies of decision making in humans using functional magnetic resonance imaging. Early examples by O'Doherty et al. (2004) used a temporal-difference (TD) learning model to estimate the reward prediction errors that subjects should have experienced and found correlates of action-independent and action-dependent learning in the ventral and dorsal striatum, respectively. Tanaka et al. (2004) took a set of TD learning models with different temporal discounting parameters and found ventrodorsal maps of prediction time scales in the striatum and the insular cortex. More recently, Daw et al. (2006) investigated the neural mechanisms governing the exploration–exploitation tradeoff, a decision-making challenge common in uncertain and dynamic environments. To do so, they followed the now familiar program of first selecting a computational model among several competing possibilities based on its ability to fit a subject's overt behavior and then using the model's hidden variables to classify each of the subject's decisions as either exploratory or exploitative. This classification allowed the authors to localize brain areas whose activity was specifically correlated with exploratory behavior.

What is required of intermediary behavioral models

There are two essential requirements for a model to be useful in this capacity. First, and most obviously, the model must contain an internal variable that operationalizes the decision variable under study. Second, the model must comprise a sufficient description of the choice behavior in the experiment. This sufficiency should be assessed according to two separate, but equally important, criteria: first, the ability of the model to predict the subject's next choice given the history of events that preceded the current trial, and second, the ability of the model to generate realistic patterns of choice in simulation, independent of the history of the subject's actions (Corrado et al., 2005).

It is worth noting the conspicuous absence of general "biological plausibility" as a requisite property of the type of behavioral model that we are proposing. The essential function of the model here is not necessarily to serve as an explicit hypothesis for how the brain makes a decision, but only to formalize an intermediate decision variable. Thus, even if a model that contains some components that seem unlikely to be literally implemented in the brain (e.g., infinite memories or perfect Bayesian inference), the resulting proxy variables may prove useful if they correctly capture the essential concept that the researcher wishes to study.

Conclusion

The generality of this new approach and the success that it has enjoyed to date bode well for its continued use in cognitive neuroscience, within the field of decision making and beyond. However, in designing future research programs around this methodology, we should be mindful of certain caveats. Detecting neural activity correlated with an internal variable of a particular model

does not necessarily mean that the brain literally implements that same computational algorithm. Finding such correlates does, however, help us to focus on a particular area of the brain or type of neuron in guiding decisions. Ultimately, to clarify the actual computational mechanisms at play and to establish causal relationships with behavior, this approach should be complemented by rigorous, falsifiable hypothesis testing and possibly intrusive manipulations of neural function.

References

- Andersen RA (1995) Encoding of intention and spatial location in the posterior parietal cortex. *Cereb Cortex* 5:457–469.
- Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA (1996) A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis Neurosci* 13:87–100.
- Colby CL, Goldberg ME (1999) Space and attention in parietal cortex. *Annu Rev Neurosci* 23:319–349.
- Corrado GS, Sugrue LP, Seung HS, Newsome WT (2005) Linear-nonlinear-Poisson models of primate choice dynamics. *J Exp Anal Behav* 84:581–617.
- Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. *Nature* 441:876–879.
- Dorris MC, Glimcher PW (2004) Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44:365–378.
- Doya K (2007) Reinforcement learning: computational theory and biological principles. *HFSP J* 1:30–40.
- Herrnstein RJ (1961) Relative and absolute strength of responses as a function of frequency of reinforcement. *J Exp Anal Behav* 4:267–272.
- Hubel DH, Wiesel TN (1959) Receptive fields of single neurones in the cat's striate cortex. *J Physiol (Lond)* 148:574–591.
- Huk AC, Shadlen MN (2005) Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *J Neurosci* 25:10420–10436.
- Lau B, Glimcher PW (2005) Dynamic response-by-response models of matching behavior in rhesus monkeys. *J Exp Anal Behav* 84:555–579.
- Nash J (1950) Equilibrium points in n-person games. *Proc Natl Acad Sci USA* 36:48–49.
- Newsome WT, Britten KH, Movshon JA (1989) Neuronal correlates of a perceptual decision. *Nature* 341:52–54.
- O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304:452–454.
- Padoa-Schioppa C, Assad JA (2006) Neurons in the orbitofrontal cortex encode economic value. *Nature* 441:223–226.
- Platt ML, Glimcher PW (1999) Neural correlates of decision variables in parietal cortex. *Nature* 400:233–238.
- Samejima K, Doya K, Ueda K, Kimura M (2004) Estimating internal variables and parameters of a learning agent by a particle filter. *Adv Neural Info Processing Sys* 16:1335–1342.
- Samejima K, Ueda Y, Doya K, Kimura M (2005) Representation of action-specific reward values in the striatum. *Science* 310:1337–1340.
- Shadlen MN, Newsome WT (1996) Motion perception: seeing and deciding. *Proc Natl Acad Sci USA* 93:628–633.
- Sugrue LP, Corrado GS, Newsome WT (2004) Matching behavior and the representation of value in the parietal cortex. *Science* 304:1782–1787.
- Sugrue LP, Corrado GS, Newsome WT (2005) Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat Rev Neurosci* 6:363–375.
- Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S (2004) Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat Neurosci* 7:887–893.
- Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. *Nature* 398:704–708.