Designing and Interpreting Psychophysical Investigations of Cognition

Michael L. Waskom,[1] Gouki Okazawa,[1] and Roozbeh Kiani[1,2,3,*]
[1]Center for Neural Science, New York University, 4 Washington Place, New York, NY 10003, USA
[2]Neuroscience Institute, NYU Langone Medical Center, 550 First Avenue, New York, NY 10016, USA
[3]Department of Psychology, New York University, 4 Washington Place, New York, NY 10003, USA
*Correspondence: roozbeh@nyu.edu
https://doi.org/10.1016/j.neuron.2019.09.016

Scientific experimentation depends on the artificial control of natural phenomena. The inaccessibility of cognitive processes to direct manipulation can make such control difficult to realize. Here, we discuss approaches for overcoming this challenge. We advocate the incorporation of experimental techniques from sensory psychophysics into the study of cognitive processes such as decision making and executive control. These techniques include the use of simple parameterized stimuli to precisely manipulate available information and computational models to jointly quantify behavior and neural responses. We illustrate the potential for such techniques to drive theoretical development, and we examine important practical details of how to conduct controlled experiments when using them. Finally, we highlight principles guiding the use of computational models in studying the neural basis of cognition.

The overarching aim of cognitive neuroscience is to understand, through experimental investigation, the mechanisms that give rise to intelligent behavior. Progress depends on many factors, with experimental design quality playing a central role. Effective experimental designs will induce controlled alterations in cognitive processes that can be related to changes in behavior and neural responses. The precision of the control determines the clarity of the theoretical insights that can be gained.

Experimental control is multi-faceted (Boring, 1954). Broadly, the objective is to produce a known change in some component of a system without directly altering any of its other aspects. When successful—and combined with an understanding of the system's overall functional goal (Krakauer et al., 2017)—this will license conclusions about how component operations within the system give rise to its emergent properties. That is, it will provide the basis for a mechanistic explanation (Bechtel, 2007). Conclusions about mechanism are limited when there is uncertainty regarding the character or magnitude of the change produced by a manipulation or in the presence of experimental confounds: failures of control that allow the values of other variables to change along with the component of interest.

Achieving precise control when investigating the neural basis of cognition can be fraught with difficulty. Cognitive processes are influenced by sensory inputs, and they are ultimately realized in behavioral responses. Yet they are situated at a far remove from the external variables that an experimenter can manipulate and measure. The intervening systems are complex, parallel, and interactive. At best, they are only partially understood. Without great care, there will be considerable uncertainty about the effects of an experimental manipulation on the cognitive process of interest, and confounding changes in other processes will concomitantly occur (Friston et al., 1996). Experimental remoteness therefore poses a fundamental challenge (Figure 1).

Neural recordings or interventions on neural activity may seem more proximal to internal processes. But clear interpretation of such data depends just as strongly on experimental control as do purely behavioral experiments. Further, neural activity is not directly interpretable in terms of cognitive processing without a theoretical framework that can bridge across cognitive and neural levels of analysis (Farrell and Lewandowsky, 2018; Marr, 1982; Teller, 1984). Such frameworks can be formally instantiated in computational models, but the benefits of doing so are limited if an experiment does not permit quantification of behavioral and neural measurements.

This review discusses principles of experimental design and interpretation that can help to overcome the challenge of experimental remoteness. Our primary goal is to share insights from an approach that applies psychophysical techniques to investigate the mechanistic basis of cognition. We first introduce this cognitive psychophysics research program and demonstrate its potential for making progress on understanding cognitive processes in experiments that prioritize experimental control. We then identify important practical issues related to design and interpretation that arise in these experiments. We further illustrate the challenge of experimental remoteness and the utility of our proposed approach by examining a pervasive issue in cognitive neuroscience: the potential for uncontrolled task engagement confounds. Finally, we conclude with a brief discussion of how to develop and evaluate computational models. In raising theoretical and practical challenges, we share specific answers from the perspective of cognitive psychophysics, but we also identify general principles that, if applied in any research program, would help to ensure a successful progression of scientific insights.

Reducing Experimental Remoteness through Cognitive Psychophysics

Sensory psychophysics has a long history of using quantitative decision theory as a tool to aid the investigation of perceptual systems (Green and Swets, 1966; Link, 1992). The cognitive
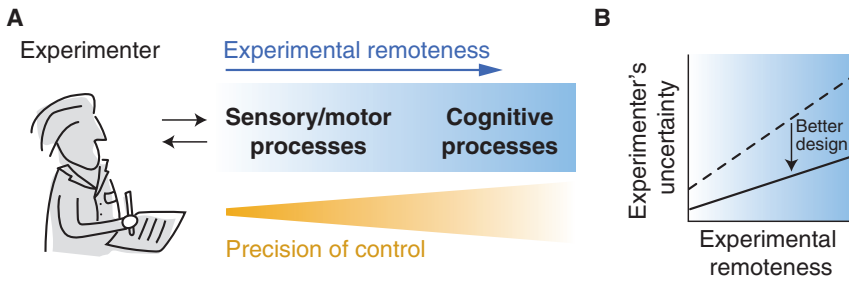
Figure 1. The Challenge of Experimental Remoteness
(A) Cognitive processes are less accessible to manipulation and measurement (experimental remoteness), posing difficulties for experimental control.
(B) In general, an experimenter's uncertainty about manipulation, measurement, and interpretation increases with remoteness. However, good experimental design choices can reduce the strength of this relationship.

psychophysics research program inverts this basic logic, using experimental paradigms and models from sensory research as tools to aid the study of cognitive processes. Doing so involves several important and mutually constitutive elements. The first is a focus on tasks where expert observers make threshold-level judgments about experimental stimuli that afford precise control. The second is a conceptual orientation toward quantification as a key goal of experimental design. The third is the centrality of formal computational models to the analysis and interpretation of behavioral and neural data. Individually, these components may not be unique to psychophysical research. Together, they allow one to approach questions about cognition from a position that remains deeply rooted in sensory psychophysics. Existing knowledge about perceptual systems then forms a bridge that can reduce experimental remoteness.

This approach has been applied successfully within the domain of perceptual decision making, often by adopting perceptual discrimination tasks directly from sensory psychophysics. Such tasks typically involve synthetic, parameterized stimuli (Rust and Movshon, 2005). For example, subjects might be asked to discriminate the relative frequency of two tactile vibrations (Hernández et al., 2000), the dominant odor in a mixture (Kepecs et al., 2008), or the direction of coherent motion in a random dot kinematogram (Shadlen and Newsome, 2001). The advantage of synthetic stimuli is that the information available to the subject can be precisely modulated along a continuous dimension, often with ratio scaling (Fechner et al., 1966). And efforts to model decision-making processes can reflect not just knowledge of the stimulus parameters themselves but also the properties of the responses that those stimuli elicit in sensory cortex (Shadlen et al., 2006).

These elements have contributed to a basic understanding of the mechanisms that underlie decision making in simple discrimination tasks. Central to this understanding is the notion of a gradual accumulation—formally, integration with respect to time—of sensory "evidence," producing a "decision variable" representation. The value of the decision variable can be used to select a response, and its computation appears to be reflected in the dynamics of trial-averaged firing rates recorded from neurons in multiple cortical and subcortical areas (Gold and Shadlen, 2007; Hanks and Summerfield, 2017; Schall, 2001). When aligned to the time of the response in a reaction time task, activity in these neurons converges onto a terminal rate (Roitman and Shadlen, 2002). This is consistent with the computational idea that a single mechanism—a decision bound or threshold—can explain both choice and decision time. The

position of the bound determines the relative trade-off between the expected speed and accuracy of the decision (Ratcliff and Rouder, 1998). These observations demonstrate that the processes studied in simple discrimination tasks are both deliberative and flexible, bearing hallmarks of higher cognition.

Ongoing work continues to provide a more sophisticated understanding of these basic elements while also exposing current gaps in knowledge (Hanks and Summerfield, 2017; Najafi and Churchland, 2018). We focus here on two strands of work that illustrate the potential of the cognitive psychophysics approach. The first is the development of expanded perceptual judgment tasks and their use in quantifying the properties of cognitive systems that support deliberative reasoning. The second shows how understanding the neural mechanisms of the speed-accuracy trade-off provides a new view on a long-standing question in cognitive control research.

Befitting their origins in the study of perception, classical tasks strongly associate the decision reported by the subject with a unified percept of the stimulus. Recent work in cognitive psychophysics has begun to develop tasks that solicit judgments about sequences or ensembles of stimuli, breaking the singular association between perception and decision while retaining the benefits of simple parameterization (Figure 2A). For example, subjects may be asked to discriminate the relative rate of discrete auditory clicks or visual flashes (Brunton et al., 2013; Raposo et al., 2014; Scott et al., 2017) or to infer something about the generating statistics of the contrast, orientation or other properties in a sequence of gratings (Cheadle et al., 2014; Drugowitsch et al., 2016; Waskom and Kiani, 2018).

These tasks have taken questions about the cognitive systems involved in deliberative reasoning and made them accessible to psychophysical quantification of the factors that limit threshold-level performance. For example, Brunton et al. (2013) leveraged the precise timing of stimulus information in a "Poisson clicks" task to determine that stochastic variability in choice is stimulus-dependent rather than time-dependent, implying that the accumulation process itself is virtually noiseless (Figure 2B). While this might suggest that expanded perceptual judgment tasks are still just studying sensory processes, Drugowitsch et al. (2016) used an orientation judgment task to show that sensory noise per se cannot account for the pattern of errors that humans make. Instead, the noise that limits behavior emerges during the cognitive process that transforms sensory representations into a common currency of decision evidence; that is, during inference (Figure 2C). Waskom and Kiani (2018) combined elements from each of these tasks to quantify the
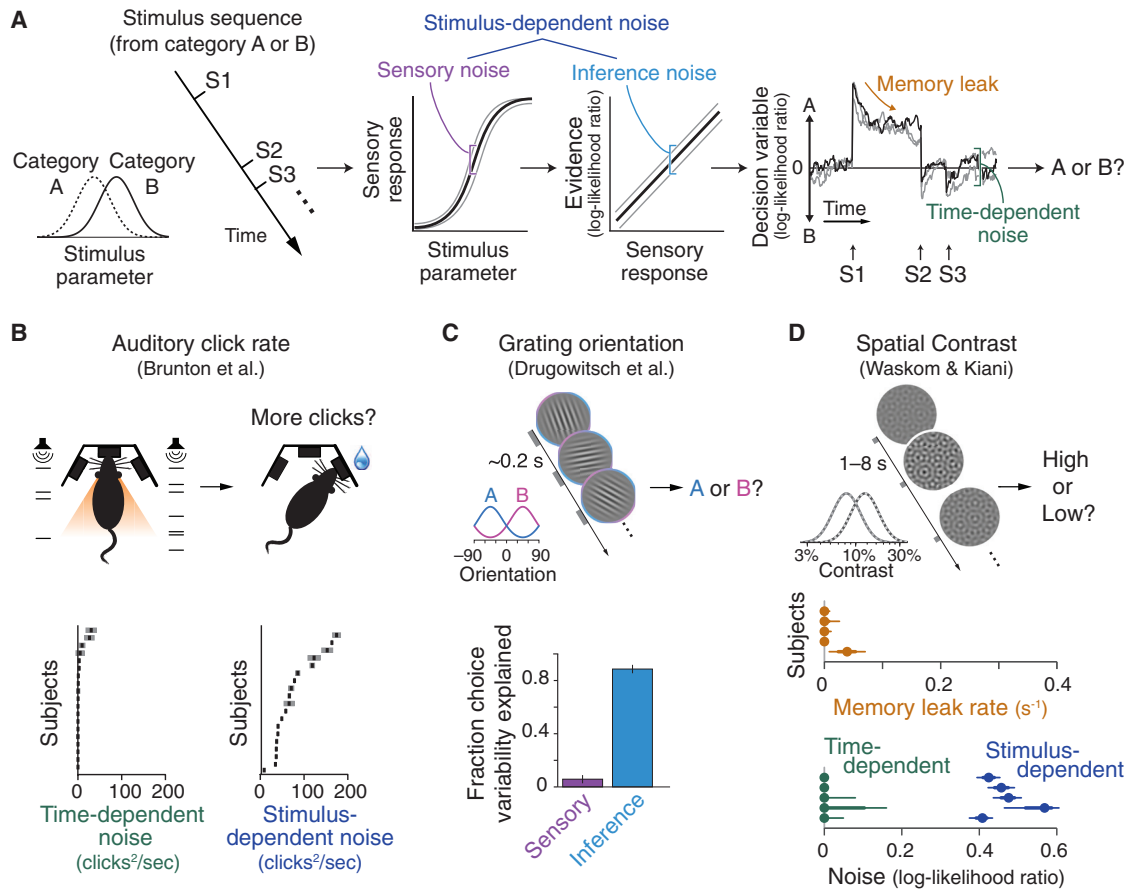
Figure 2. Expanded Perceptual Judgment Tasks Enable Quantification of Deliberative Reasoning Behavior
(A) The components of an expanded perceptual judgment task. Subjects make inferences about the latent state of the world using multiple pieces of unreliable sensory information (i.e., "are these stimuli drawn from category A or B"?). Stimulus strength influences the internal sensory response, which must be converted into a representation of evidence that bears on the specific inference problem. Stimulus-dependent noise can be introduced at either stage of processing. Normative inference can be achieved by integrating information from multiple stimuli, but this process may be limited by time-dependent factors such as leak or noise in memory.
(B) A rate discrimination task using auditory "clicks." Behavioral quantification shows that the source of internal noise is dependent on the appearance of stimuli rather than on the passage of time. From Brunton et al. (2013).
(C) An orientation judgment task with high-contrast gratings. Separately quantifying sensory and inferential contributions to choice variability suggests that stimulus-dependent noise arises primarily from cognitive systems. From Drugowitsch et al. (2016).
(D) A contrast judgment task where long (1–8 s) gaps separated each appearance of a hyperplaid stimulus. Minimal influence of time-dependent limitations reveals that evidence integration can make use of a robust memory system. From Waskom and Kiani (2018).

mnemonic properties of deliberation. By manipulating the times at which spatial contrast patches were presented, they confirmed that time-dependent noise is minimal and quantified the integration time constant of behavior as being on the order of tens of seconds or longer. This implies that deliberative reasoning can make use of a memory system that has considerable robustness in the temporal domain (Figure 2D).

Developing and evaluating the computational models that contributed to these insights required precise knowledge about the timing and strength of the information available for the decision. Simple parameterized stimuli provide the necessary precision. Knowledge of how those stimuli are represented in sensory systems can aid interpretation by accounting for nuisance factors that might otherwise complicate the analysis. For example, Brunton et al. (2013) included auditory adaption in their model, and Drugowitsch et al. (2016) accounted for systematic biases

in the representation of near-cardinal orientations, each helping to refine the quantification of cognitive factors that were the target of the investigation.

An important characteristic of the models used for research on perceptual decision making is that they link the parameters of each stimulus to the speed and accuracy of the corresponding perceptual judgment. That is, they are models of behavior (Krakauer et al., 2017). The results highlighted above emphasize that these models are not limited to quantifying the properties of sensory representations: they can also inform our understanding of the cognitive processes that use those representations to reason about the world (Bogacz et al., 2006; Gold and Stocker, 2017; Link, 1992; Shadlen and Kiani, 2013). Therefore, they can be a tool for reducing experimental remoteness. Because evidence accumulation models portray behavioral responses as the product of a dynamical process, they can also provide a formal guide
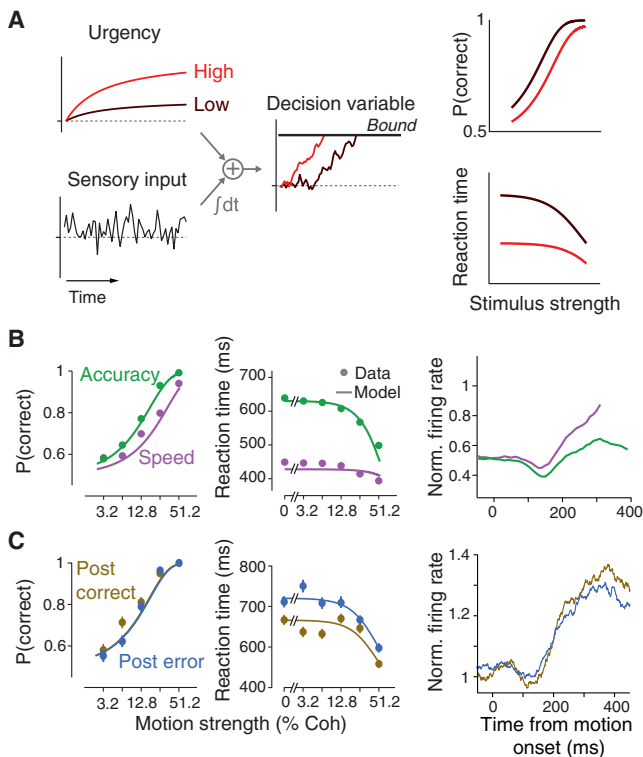
Figure 3. Models of Perceptual Discrimination Suggest Mechanisms for Cognitive Control

(A) Schematic illustrating a class of computational models that attribute trade-offs between speed and accuracy to the strength of an evidence-independent "urgency" signal. The urgency signal exerts an additive influence on the decision variable, pushing it toward a termination bound. When urgency is high, decisions are made faster, but they are less likely to be correct. The urgency signal is a potential mechanism for higher-order control processes to act through.

(B) Behavioral and neural data from an experiment where monkeys made perceptual discriminations while prioritizing either speed or accuracy. The model fits to behavior in the first two panels were constrained by deriving an urgency signal from neural activity, which is shown in the final panel. From Hanks et al. (2014).

(C) Decisions following an error are slower but no more accurate. Behavioral quantification shows that this post-error slowing can be explained by an error-dependent reduction in sensory sensitivity along with a compensatory decrease in urgency. From Purcell and Kiani (2016).

for interpreting the dynamics of neural responses in decision-making tasks. This is key, because understanding neural activity in terms of the computations that it implements is a prerequisite for developing a mechanistic explanation of complex behavior (Bechtel, 2007; Carandini, 2012; Krakauer et al., 2017).

Explaining neural responses in terms of their computational parameters also makes it possible to understand how those parameters are modulated by higher-order control systems. An illustrative example concerns strategic adjustments in the trade-off between speed and accuracy. As previously mentioned, the common pre-saccadic firing rate for decision-related responses evokes the concept of a terminal threshold or bound, suggesting that an emphasis on speed or accuracy would lead to changes in its amplitude. Surprisingly, this turns out not to be the case (Figure 3A). Instead, an emphasis on accuracy decreases the strength of an evidence-independent

"urgency" signal, leading to a lower starting point and shallower rate of growth (Figure 3B) (Hanks et al., 2014). This unexpected result might have been dismissed or misinterpreted had the neurally derived urgency signal not quantitatively accounted for changes in behavior.

Experimental subjects can be encouraged to adopt different speed-accuracy policies either through direct instruction (Palmer et al., 2005; Ratcliff and Rouder, 1998) or by manipulating the temporal statistics of a task (Hanks et al., 2014; Heitz and Schall, 2012). Cognitive control processes also appear to enact endogenous adjustments in this balance. One such adjustment is evident when reaction times increase immediately following errors, a phenomenon known as "post-error slowing" (Rabbitt and Rodgers, 1977). A persistent challenge in understanding post-error slowing has been the rarity of observing corresponding improvements in accuracy, as would be expected by a simple application of the logic underlying the speed-accuracy trade-off (Danielmeier and Ullsperger, 2011). Purcell and Kiani (2016) addressed this question using model-based analyses of a simple perceptual discrimination task. They found that slower responses immediately following an error reflect an adaptive compensation for a temporary decrement in perceptual sensitivity. Both behavioral modeling and analyses of neural responses indicated that this compensation was accomplished via changes in the urgency signal's amplitude (Figure 3C).

The goal of the preceding discussion has been to illustrate the potential of the cognitive psychophysics research program. The examples demonstrate that it is possible to gain understanding of cognitive mechanisms by taking an "outside in" approach that incrementally progresses from peripheral to central systems. Insights about sensory representation lead to models of the process underlying deliberative choice in discrimination tasks, which themselves suggest mechanisms for executive control. Each later stage leverages earlier insights for improved experimental design and interpretation. Despite this success, we do not mean to imply that there are no open questions or controversies relating to these topics. Indeed, there has been and will continue to be debate about both the structure of the computational models (Tsetsos et al., 2012; Drugowitsch et al., 2012; Hawkins et al., 2015; Miller and Katz, 2013; Moran, 2015; Ratcliff et al., 2016; Thura and Cisek, 2016) and the neural mechanisms that implement these computations (Scott et al., 2017; Huk et al., 2017; Servant et al., 2019). Yet even where controversy remains, the formal models elevate the process of reconciliation above a terminological dispute.

There are also many important questions in cognitive neuroscience that are unlikely to be answered through further elaboration of the perceptual decision-making paradigm, and we would not suggest that aiming to do so is the best path forward in every case. The work described above was able to mitigate experimental remoteness through an intense focus on experimental control and by taking advantage of accumulated knowledge about peripheral sensory systems and intervening computations. A similar outside-in approach has been successfully pursued elsewhere. For example, recent computational insights into working memory have been enabled by a similar focus on simple tasks that permit incorporation of knowledge about sensory representations into models of capacity limitations

(Ma et al., 2014). Working from the other end of the system, the field of computational motor control has grown from studying the basic psychophysics of simple movements toward a quantitative account of cognitive operations such as learning, planning, and decision-making in skilled sensorimotor behavior (Gallivan et al., 2018; McDougle et al., 2016; Schall, 2001; Wolpert et al., 2011). We are optimistic that the cognitive psychophysics approach can be successfully pursued in other domains as well.

Investigations of executive functions that are rooted in either sensory or motor psychophysics would have particular advantages in both experimental control and interpretational clarity. The reason is that they can study the influence of higher-order processes on systems that are reasonably well understood from both a computational and physiological perspective. To motivate this idea, we observe that cognitive control is commonly thought of as involving a goal-directed parameterization of sensory, cognitive, or motor systems (Miller and Cohen, 2001). An outside-in approach begins with an understanding of the parameters that the control system should set, along with suggestions about how it might do so. The preceding discussion of the speed-accuracy trade-off illustrates this point: the computational construct of a decision bound generated clear predictions about how to model adaptive adjustments of decision policy. And the knowledge that changes in the bound height could manifest at the neural level through an urgency signal suggested a clear hypothesis about how those adaptive adjustments might be implemented in the brain.

In advocating for an outside-in approach, we do not envision a completely serial progression of research topics. The advantages of the approach do not require full understanding of each stage of information processing prior to or following those at which cognitive systems exert their influence, nor do the resulting models need to include the full complexity of existing knowledge about sensory or motor systems. Indeed, an important benefit of working with first-order systems that are reasonably well-understood is a freedom to abstract away variables that are known to be unimportant. But the challenge of experimental remoteness will limit mechanistic insight when investigations into higher-level cognition are pursued without clear understanding of the lower-level processes that are closer to the variables one directly manipulates and measures.

Elements of Psychophysical Investigation

Having highlighted some of the scientific insights contributed by the cognitive psychophysics research program, we transition now to a discussion of more practical details that arise during these investigations. We aim to highlight some key design choices that are important for achieving strong experimental control, especially ones that may not be obvious when reading papers that focus on novel theoretical advances. As before, most of our specific examples will be drawn from experiments on perceptual decision making, but many of the issues that we raise reflect general challenges that any investigation of cognition faces. Therefore, we have tried to extract general principles from our specific examples. These principles would be useful to incorporate more broadly, wherever possible.

To make best use the precise control offered by sensory stimuli, one needs to choose wisely when picking specific stimulus values to show. A best practice is to cover a broad range and to sample densely along it. Ideally, stimuli will range from completely non-informative—such as a random dot kinematogram with 0% coherence or two trains of auditory clicks with identical rates—to values that an attentive subject should be able to categorize correctly on every trial. Stimuli sampled between these extremes will vary in strength, and they should do so smoothly enough so as not to appear in discrete categories of difficulty. At intermediate values, stimuli will span a "threshold" where performance is limited by the properties of the sensory and cognitive systems that are used to perform the task. Explaining the factors that lead to a particular value of the threshold is a key goal of psychophysics (Green and Swets, 1966; Parker and Newsome, 1998; Quick, 1974), but embracing the full range is critical for strong experimental design. There are two reasons for this.

The first reason is that successfully predicting the particular shape of the psychometric function across a broad range of densely sampled stimuli poses a high bar for candidate models to clear. Consider the differences between two classes of models. One class can predict the accuracy of choices about any stimulus between two broad extremes, along with predicting the accompanying reaction times and subjective confidence. The other can make only ordinal predictions about individual variables, such as that subjects would be more accurate when judging "easy" stimuli compared to "hard" stimuli. As a general principle, models are more powerful when they make more specific predictions, especially about multiple facets of behavior (Farrell and Lewandowsky, 2018; Smith and Little, 2018). But this will be true only if the experimental design generates rich enough data to properly evaluate the more sophisticated models.

The second reason has less to do with one's theory of the sensory or cognitive process itself and more to do with characterizing how the subject is approaching the behavioral task: that is, the subject's strategy. Staircase methods, by focusing on a specific level of performance, can efficiently estimate values of the subject's psychophysical threshold (Cornsweet, 1962). Nevertheless, they reduce an experimenter's ability to uniquely interpret the cause of errors. With designs that sample stimuli randomly and more broadly—ranging from chance performance to perfect accuracy—simple forced-choice data can provide rich information about confounding factors such as undesired strategies or inconsistent task engagement. Stimuli along the full range of the psychometric function can aid this level of understanding, albeit in different ways.

Even when subjects should be able to judge the easiest stimuli correctly on every trial, they may not always do so. It is common to refer to the deviation between asymptotic performance and perfect accuracy as the subject's "lapse rate" (Kingdom and Prins, 2010). The terminology is drawn from a typical explanation for this effect as reflecting momentary "lapses" in attention. An attentional lapse might cause a subject either to miss the stimulus and then guess randomly or to accidentally make a motor response that does not reflect their subjective judgment. As a practical matter, the lapse rate corresponds to the asymptote of the psychometric function, not just to performance in the highest stimulus strength condition. However, to estimate this value

well, the experimental design must sample appropriately extreme stimulus values. Doing so is important because frequent lapses can indicate more fundamental problems in an experiment. At best, occasional task disengagement might be thought of as adding random, unbiased noise to a dataset. Modest additional measurement noise may seem tolerable, but it is worth considering whether you would continue to use a response collection device that records random data on 10% of trials, as implied by 95% accuracy on the easiest trials (Fetsch, 2016).

Significant lapse rates in a perceptual discrimination task may also imply that the subject is selecting actions on the basis of factors other than the stimulus. For instance, subjects may be influenced by the recent history of choices and outcomes, even when trials are independent (Abrahamyan et al., 2016; Akrami et al., 2018). Alternatively, lapse rates may reflect exploratory responses made when subjects are uncertain of the task structure (Pisupati et al., 2019) or otherwise seeking information (Sugrue et al., 2005). Whether or not these factors pose a challenge for interpretation depends entirely on the goals of the experiment. In an experiment designed to investigate stimulus-guided behavior, exploratory or history-based responses represent confounding factors that will reduce interpretability. Yet, in many natural environments, recent history is informative, and exploration is adaptive (Glimcher, 2003). Therefore, understanding how the decision-making process incorporates these factors is an important goal. Ideally, this effort would not just account for them in models but also bring them under experimental control and devise conditions where they can be eliminated. Doing so would represent a generalization of the logic behind lapse rate analysis: eliminating the influence of a process is a strong demonstration of the control that one has over it.

At the other end of the spectrum, perceptual discrimination experiments may also include trials where the stimulus is positioned exactly at the category boundary. Why include trials where there is, by definition, no correct answer? When characterizing a system, it can be useful to put in noise and see what comes out (Marmarelis and Naka, 1972). For example, non-informative stimuli will help to characterize directional or history-dependent response biases. More elaborate methods that leverage non-informative stimuli include psychophysical reverse correlation, which computes a choice-conditioned, temporally resolved average of stimulus intensities across trials. Reverse correlation estimates the "psychophysical kernel," or weighting of the stimulus across time. The psychophysical kernel allows one to infer the internal dynamics of sensory and decision processes from simple forced-choice behavior (Neri and Heeger, 2002; Nienborg and Cumming, 2009; Okazawa et al., 2018).

Sometimes, the shape of the psychophysical kernel bears directly on the main research question. This is the case when it is used as an assay of primacy or recency biases, which cause information presented earlier or later in the trial to exert a relatively larger influence on the decision (Cheadle et al., 2014; Kiani et al., 2008; Tsetsos et al., 2012). Even when such biases are not the focus of an experiment, however, determining whether or not they are present can be important for justifying assumptions about decision strategy. For instance, signal detection theoretic and drift diffusion models assume perfect integration across the duration of a stimulus. Departures from perfect integration,

which can be identified using reverse correlation, can confound estimates of behavioral or neural sensitivity using these models (Okazawa et al., 2018). Such confounds can be removed post hoc only if one knows the mechanisms that give rise to primacy or recency and can incorporate them during modeling. Neither requirement is trivial.

More generally, any model-based investigation depends on the validity of the assumptions underlying the model. This is broadly understood in the context of assumptions about statistical distributions, but investigating the neural mechanisms of cognition also entails assumptions about cognitive processes themselves. Finding ways to justify those assumptions through a characterization of the subject's behavior is a critical part of experimental design and interpretation. This can be particularly challenging when the assumptions involve cognitive processes that lie outside the scope of the modeling framework itself. The previous example illustrates such a case, where analyses of perceptual sensitivity could be confounded by poor understanding of decision strategy, a more abstract (and less-well understood) construct. Behavioral characterization must therefore be more thorough than simply testing the hypothesis of interest: it should also validate assumptions about other cognitive processes that are implied by one's computational framework.

While identifying deviations from assumed task strategies is important, care must also be taken to avoid directly confounding task manipulations with changes in strategy. In the context of a perceptual discrimination experiment, this confound might arise if stimulus strengths are sampled too coarsely. Standard models assume that strategy is invariant to stimulus strength. This is a reasonable assumption when different difficulty conditions are randomly interleaved and treated as latent variables, known to the experimenter but not explicitly signaled to the subject. Conditions that are blocked or that differ too obviously in difficulty, however, may cause a subject to adopt a mixture of strategies. This could be a relatively subtle confound, emerging as a condition-dependent speed-accuracy trade-off: a subject might wait to set the position of their decision bound until they have an initial sense about the reliability of a stimulus (Figure 4B). More dramatically, subjects might adopt entirely different computational solutions for the different conditions. For example, they might accumulate evidence only when faced with a weak stimulus, solving the task on easier trials through "snapshot" or "extrema detection" strategies, where choices are made based on single observations that are selected either randomly or when one sample exceeds a large decision criterion (Quick, 1974; G. Stine et al., 2018, Cosyne, conference; Waskom and Kiani, 2018).

When parametric manipulations are implemented through explicit instructions to the subject, rather than by changing the latent parameters of a stimulus or task, it can elevate the probability of a confound with strategy. A classic example arises in the *n*-back working memory task, where subjects view a stream of stimuli and must indicate when the current stimulus matches the one presented *n* trials prior. In principle, parametrically manipulating *n* allows one to alter the load on working memory along a continuous dimension. Bringing the logic of parametric manipulation to the cognitive neuroscience of working memory (Braver et al., 1997) represented an early advance beyond simple subtraction designs (Friston et al., 1996). But direct
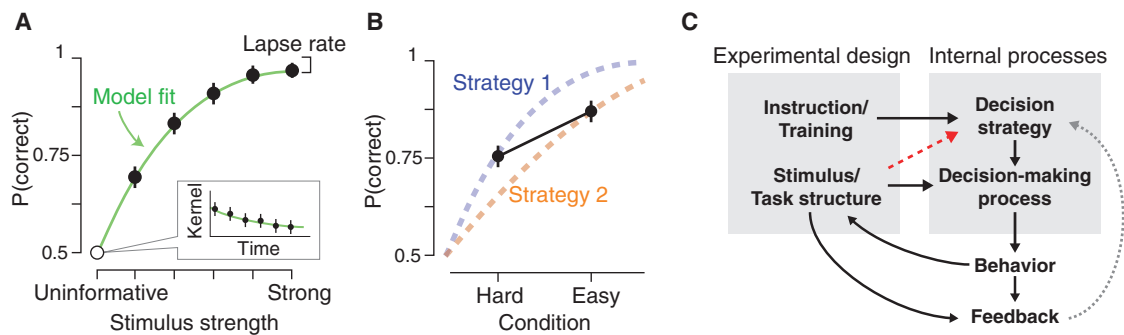
Figure 4. Experimental Design Can Both Constrain and Reveal Decision Strategy
(A) Schematic illustrating how the psychometric function affords rich characterization of behavior. Densely sampling stimulus strengths produces a narrow target for model development. Strong stimuli can be used to estimate the lapse rate, while uninformative stimuli can be used to estimate the psychophysical kernel.
(B) Schematic illustrating how sparse sampling of conditions can produce mistaken interpretations. The dashed lines indicate psychometric functions that arise from condition-dependent strategies. For example, the subject's speed-accuracy trade-off might vary for easy and difficult stimuli if they are the only stimuli used in the experiment. Sparsely sampled designs are limited in their ability to distinguish different factors that influence behavior, and they may even induce confounds.
(C) Control involves both enforcing and avoiding dependencies between the experimental design and internal processes. When investigating the mechanisms of decision making, it is best if task instructions and training alone determine the subject's strategy. Strategy should not vary along with other experimental conditions, but poor design choices might introduce such a dependency (red arrow). Strategy can also change from trial to trial depending on the outcome of each choice (gray arrow). Whether or not this is a confound depends on the goals of the experiment.

instruction of the parametric variable makes it possible for subjects to adopt different strategies at each level. For example, a subject may comfortably perform a 1-back or even 2-back task by maintaining each item in working memory but resort to using familiarity-based judgments—a long-term memory mechanism—when their capacity is exceeded (Kane et al., 2007).

Confounds that involve condition-dependent changes in strategy pose serious challenges for efforts to quantify cognitive processes or interpret neural responses. As an example, imagine attempting to test a model of how choice confidence is represented in the brain by using only two levels of stimulus difficulty. If the subject's strategy changes between conditions, it would be impossible to attribute any differential neural activity to confidence per se rather than to those different strategies. When stimuli can be parametrically controlled, there is a relatively simple solution to alleviate the problem: sample stimulus strengths more densely. Dense sampling increases uncertainty about stimulus difficulty and reduces the likelihood that subjects will adopt stimulus-dependent strategies. As a general rule, the more subtle the experimental manipulation, the more difficult it will be for a subject to adopt condition-dependent strategies.

It is also advisable to isolate cognitive variables through multiple orthogonal manipulations. For instance, modulating the duration of a perceptual stimulus will influence choice accuracy (and confidence) independently of its strength (Khalvati and Rao, 2015; Kiani and Shadlen, 2009). Stimulus duration manipulations are less likely to introduce strategy confounds, but their influence on confidence is also weaker. Implementing both manipulations can be a way to achieve a balance between power and control. Aiming for such a balance can be a useful general goal. Ideally, the effects of multiple manipulations will be explained within a unified computational framework.

As the previous point emphasizes, control over the temporal aspects of a task can have important consequences for cognitive processes. Under the assumption that subjects employ some form of accumulation or sequential sampling, stimulus duration influences the amount of information available for the choice. But showing a stimulus for a long duration does not guarantee that the subject will use all of that information. Actively engaging with and accumulating sensory information is costly (Drugowitsch et al., 2012), encouraging subjects to satisfice by using only partial information (Kiani et al., 2008). With long, fixed duration trials, it may not be possible to know when the subject was or was not engaged. This can complicate both quantification of behavior and analyses of neural data. Variable-duration designs enhance experimental control, although the choice of duration statistics is important (Ghose and Maunsell, 2002; Nobre and van Ede, 2018). Uniform distributions prevent a subject from predicting the duration of an event before it begins, but they have an increasing hazard function, meaning that the subject can anticipate the end of the event as it progresses. Such anticipation will not be possible if durations are distributed exponentially (or geometrically, in the case of discrete stimulus presentations), but this choice implies that most trials will be relatively short.

In a reaction time task, the subject controls the duration of each trial. Yet experimenters can still exert influence through stimulus sampling. Doing so is important because, while many models make predictions about temporal aspects of cognitive processing, reaction time is not a pure measure of processing time. This is partly due to sensory and motor latencies—"non-decision time" in the parlance of drift diffusion models—and also because subjects may vacillate between choices or hesitate before responding. The advice about sampling stimulus strengths broadly is useful on both points. Increasing stimulus strengths until reaction times asymptote will help to estimate the duration and variability of non-decision time. And the difference in reaction times between non-informative and very weak stimuli can help to diagnose vacillation.

Experimental design traditions in perceptual and cognitive science differ in an important respect (Smith and Little, 2018). Perceptual experiments often recruit only a few subjects; these

---

Box 1. Mechanistic Insights Require Constraint

Psychophysical methods constrain behavior and neural computations in multiple ways. These range from studying forced choice judgments about synthetic stimuli to providing subjects with extensive instruction and training before data collection. Such practices, which are relatively noncontroversial for researchers who study sensation and perception, may strike those who work in other areas as unusual. Are we not missing something by studying highly constrained and arguably artificial behaviors?

Our answer is that we are almost certainly missing things, but those things are not answers to the questions that we are asking. Building models of computational processes and their neural mechanisms requires us to treat many interesting aspects of cognition as sources of error variance that must be controlled. For example, humans often adopt heuristic strategies to avoid reasoning their way through a difficult problem (Tversky and Kahneman, 1974). Describing the heuristics that people tend to use in the real world is an important goal, as is identifying the situations in which they are likely to do so. But if our aim is to study the computations that underlie reasoning, it is essential that our experiment not be one of those situations! Descriptive investigations that do not constrain computational solutions are highly valuable; they should certainly be pursued, and it is advisable to do so with a broad scope. When one wishes to infer mechanism from limited experimental data, however, constraint is paramount.

---

subjects are highly trained, and each contribute a large number of observations. In contrast, cognitive experiments typically recruit larger samples, train for task comprehension rather than for expertise, and focus analyses on population-level parameters. The cognitive psychophysics work discussed in the previous section has largely taken the first approach (Box 1). An investigation that emphasizes quantification is most effective when subjects consistently perform at threshold, such that their behavior reflects the properties of the computational system engaged by the task rather than their strategy or level of engagement. This requires training. And many techniques for behavioral characterization must be applied at the level of the individual subject, because the adoption of undesired strategies or inconsistent levels of task engagement are likely to be expressed idiosyncratically across a population.

At the same time, the perceptual science approach is limited in its ability to quantify population level parameters or the (co)variation of individual differences, as doing so requires larger sample sizes at the subject level. This is unfortunate because, in bridging between cognitive phenomena and neural mechanisms, the models used in cognitive psychophysics would have much to contribute to the growing field of computational psychiatry (Wang and Krystal, 2014). Therefore, a focus on individual-level performance from expert subjects remains important to quantitative experimental control, but relaxing this limitation should be considered an important goal. How could our approach be scaled up to larger samples and populations that can provide only moderate amounts of data per subject? Experimental tasks that are more intuitive and naturally engaging might require less training, but developing tasks that gain these features without sacrificing the control of synthetic stimuli will not be trivial. Some of the techniques for behavioral characterization could be integrated with unsupervised learning methods to identify sub-groups of subjects with particular characteristics, allowing diagnosis of idiosyncratic strategies from limited individual-subject data. These are only vague ideas at the moment, but we hope that broader adoption of our approach will drive innovation on this front.

Variable Task Engagement as a Persistent Confound Risk

To perform a cognitive task, a subject must engage with it. As a consequence, any experiment that involves an explicit task will recruit higher-order processes related to cognitive engagement, whether or not those processes are the experiment's intended target. A manipulation that causes one task condition to be relatively more demanding than another might therefore become confounded with the subject's level of engagement. This poses particular challenges for investigating neural mechanisms of cognition, because responses in numerous cortical regions have been shown to correlate either positively (Cabeza and Nyberg, 2000; Corbetta and Shulman, 2002; Fedorenko et al., 2013) or negatively (Anticevic et al., 2012; Buckner et al., 2008; Raichle, 2015) with engagement. Many of these regions exhibit a high base rate of reported effects across the human fMRI literature (Poldrack, 2011; Yarkoni et al., 2011), suggesting that such confounds could be widespread. It is therefore important to understand how to identify, reason about, and control task engagement confounds. In this section, we highlight notable cases where the risk of task engagement confounds has been raised in the literature. We then argue that the cognitive psychophysics approach can help to ameliorate that risk.

Discussions of task engagement confounds have often centered on the role of processing time. Many techniques for measuring neural activity have the potential for confusing differences in the duration of a response with differences in its amplitude. This risk is particularly acute with fMRI, which measures a temporally integrated surrogate of neural activity (Boynton et al., 1996; Logothetis and Wandell, 2004). The concern is that, when subjects engage with a task for different durations across conditions, regions that implement general task-directed computations could exhibit different response amplitudes without explicitly representing anything about the variable that was manipulated. A notable example of this concern has arisen in debates about the anterior cingulate cortex and its role in cognitive control processes such as conflict detection and performance monitoring (Botvinick et al., 2001; Brown, 2011; Grinband et al., 2011; Yeung et al., 2011). The risk is not limited to fMRI experiments, however. Electrophysiological methods have finer temporal resolution, but they can also mistake differences in duration for those in amplitude, particularly if task engagement is not consistent across the window used to aggregate signals. Therefore, this confound is likely to emerge in experiments that lack behavioral control over the subject's temporal engagement with the task.

Analyses of neural data increasingly uses multivariate decoding techniques to characterize the information represented by distributed patterns of neural responses (Haynes, 2015). An advantage of this approach is that tasks can be designed to manipulate representational content rather than computational process. For example, instead of designing conditions that differ in demands on cognitive control, control can be investigated by studying representations of task rules that specify arbitrary stimulus-response associations (Bode and Haynes, 2009; Waskom and Wagner, 2017; Woolgar et al., 2011). It may appear that such experiments would be less likely to encounter task engagement confounds, but this is not necessarily the case. Even when cognitive processes are not intentionally manipulated, task demands might endogenously vary between conditions, leading to differences in engagement that are expressed idiosyncratically across subjects. These idiosyncratic differences would average out in population-level analyses of response amplitude. But decoding analyses are typically evaluated in terms of accuracy, an unsigned measure that does not benefit from counterbalancing (Todd et al., 2013; but see Woolgar et al., 2014). In effect, decoding analyses trade specificity for sensitivity, leaving them vulnerable to surprising confounds with task engagement. The risk will be reduced if the demands of a task can be quantitatively controlled at the level of individual subjects.

Model-based analyses offer another way to test more sophisticated hypotheses about the relationship between task manipulations and neural activity (Mars et al., 2012). In principle, formalizing variables that relate to engagement within a computational model can allow one to better separate them from constructs of interest. For example, cognitive demands in a value-based decision-making experiment should be highest when two options are most similar, allowing one to separate engagement from other representations of value (Glimcher, 2003). Yet subjects might construe the option values differently from the way that they are specified in the model, reintroducing a confound with engagement. This issue has prompted a recent debate about whether neural responses in a foraging context actually reflect domain-general control signals rather than the process of value comparison (Kolling et al., 2012, 2016; Shenhav et al., 2014). It reiterates the importance of rich behavioral characterization for validating the assumptions of one's modeling framework.

These examples demonstrate that potential confounds between experimental manipulations and task engagement can arise in multiple forms. Even when investigations of cognition pursue an outside-in approach, they will need to grapple with these higher-order phenomena. While they may be possible to mitigate during analysis, engagement confounds are best handled during experimental design. Conditions can be chosen to carefully separate variables, ideally using the predictions of formal models. Experimental subjects can be trained so that they understand the task and consistently employ an effective strategy. Tasks that afford better control over engagement and richer quantification of behavior can be employed to actively balance conditions for each subject. It will also be important to understand the mechanisms of task engagement more fully. Why are there such widespread correlations between neural activity and cognitive demands (Shenhav et al., 2017)? Is task engagement an important cognitive construct in its own right, such

that it should be incorporated into models? Or will a more complete mechanistic understanding explain it away? Resolving this uncertainty will be key for pursuing questions about more complex cognitive processes.

Cautionary Notes on Developing and Evaluating Computational Models

While computational models are central to our approach, care must be taken during model development and evaluation. The results of an individual experiment rarely provide unequivocal support for one model, but a clear understanding of the predictions made by different models can help tailor experimental designs so that competing models become maximally discriminable. The strongest model comparisons will focus on qualitative predictions: those that are robust to variations of parameters that the experimenter does not control or test. Invariance to uncontrolled or uninteresting model parameters is necessary for generalizing beyond specific model instances. Such generalizations are essential because limitations of existing knowledge about cognitive processes introduce significant uncertainties about model details, such as how different parameters are implemented or interact with each other (Busemeyer and Diederich, 2010).

The ability of computational models to generate quantitative predictions makes it easy to overlook these inherent uncertainties. If doing so leads to a strong dependence on uncontrolled or poorly understood parameters, one may arrive at incorrect conclusions. Commonly used techniques for model comparison, including those based on information criteria such as Akaike information criterion (AIC), Bayesian information criterion (BIC), or deviance information criterion (DIC) (Burnham and Anderson, 1998), penalize models for complexity but remain ignorant about the broader causes of uncertainty (Churchland and Kiani, 2016). A key point to emphasize is that statistical methods for model comparison evaluate specific model instances, but theoretical debates often revolve around distinctions between model classes. Therefore, the outcome of a formal comparison, even when statistically correct, may not generalize in the way needed to make theoretical progress.

Challenges also arise when accounting for novel experimental observations by extending existing models or starting *de novo*. Insights will be limited if these extensions or new models are generated arbitrarily. In a large model space, it is always possible to find features that will account for particular new experimental observations. Without clear guidelines for selecting these features, the success of one particular model instance will be only weakly informative.

Detailed guidelines for model development must be informed by expert knowledge and can vary across domains, but a few general principles are worth highlighting. First, new or extended models should explain the full set of reproducible experimental results, not just the new observations. Models that ignore (or fail to account for) past observations should be penalized. This principle may seem obvious, but it is easy to overlook in practice given that empirical studies are often highly focused on novel results. A second general principle is that model extensions should be supported by multiple aspects of the experimental results, as by explaining distinct facets of behavior (Palmer et al., 2005;

Box 2. Principles of the Cognitive Psychophysics Approach

Cognitive psychophysics. Experimental designs and stimuli that were originally developed to study perceptual systems can be useful for investigating cognition. Simple parameterized stimuli afford strong experimental control. This control facilitates an approach that prioritizes quantification and the development of formal computational models.

Outside-in approach. Multiple systems mediate between experimental manipulations and the cognitive processes that they target. Experimental control is enhanced when the intervening complexity is minimized and limited to systems where there is existing understanding. This can be achieved by grounding investigations of cognition in sensory or motor psychophysics.

Rich behavior. Behavior should be assayed from multiple perspectives across a broad range of experimental conditions. Rich behavior enhances model comparison by requiring more precise predictions, and it can provide opportunities for checking assumptions about strategy and task engagement. Behavioral assessments are more informative when made at the level of individual subjects.

Strategy assessment. Subjects must adopt a particular strategy to solve any experimental task. Conclusions about mechanism require knowledge of this strategy, or at least assurances about its consistency across conditions. Strong experimental designs will aim both to constrain the subject's strategy and to provide opportunities for its assessment through model-free analyses.

Task engagement. Subjects should maintain high levels of engagement throughout the task. Inconsistent engagement can be diagnosed by the lapse rate, but only when strong enough stimuli are sampled. Variations of engagement across experimental conditions can introduce challenging confounds, especially when interpreting neural responses.

Subtle manipulations. Making experimental manipulations subtle (from the subject's perspective) can help to avoid confounds with strategy. It is preferable to manipulate latent variables rather than task elements that are directly instructed. Dense sampling of stimulus space is one way to implement a subtle manipulation.

Orthogonal manipulations. Manipulating a cognitive variable through multiple independent channels provides a stronger guard against confounds. For example, confidence can be manipulated by changing either the strength of a stimulus or the duration of its presentation. The best models will account for both manipulations through a unified mechanism.

Model-based design. Computational models are necessary for bridging across levels of analysis. Experimental design should aim to craft conditions that maximize the discriminability of competing models. Ideally, models will be distinguished on the basis of qualitative (parameter-independent) predictions, not just quantitative fits.

Smith and Little, 2018), by jointly accounting for behavioral and neural data (Kiani and Shadlen, 2009; van Ravenzwaaij et al., 2017), or by using neural data to constrain fits to behavior (Hanks et al., 2014; Turner et al., 2018). A third principle is that new models, or their extensions, should have some motivation within a normative framework (Geisler, 2011; Griffiths et al., 2010; von Helmholtz, 1867; Jaynes, 2003). Normative models can provide a principled starting point or direction within a large modeling space, which helps to avoid arbitrary choices. Using them to motivate model development does not require the assumption that cognitive or neural processes are themselves optimal (Rahnev and Denison, 2018; Stocker, 2018; Summerfield and Tsetsos, 2015). Indeed, when made with care, comparisons to a normative model can generate valuable insights even if behavior or neural computations fall short of optimality.

Conclusions

In this Review, we have discussed the challenges of investigating cognition and suggested some approaches that can help to overcome them (Box 2). We have argued that a psychophysical approach—studying threshold-level judgments about synthetic stimuli that are parameterized on a continuous scale—can produce generalizable insights about cognition and drive progress toward understanding complex executive processes. This approach features tasks that afford rich, quantitative behavioral characterization while remaining grounded in knowledge of sensory and motor systems. When used judiciously, such tasks can help to constrain and diagnose a subject's strategy and level of engagement, avoiding confounds from uncontrolled higher-order processes. We have also emphasized the important role of computational models in design and interpretation while highlighting potential pitfalls in their use. We hope that these suggestions will be helpful beyond our own narrow domain.

Experimental control is essential for scientific progress. Yet it can easily be overshadowed by excitement about new tools for measurement, manipulation, or analysis. Fortunately, the growing concern about reproducibility has renewed interest in experimental methods. There are now widespread calls for an increased focus on transparency, sample size, and meta-analysis (Munafò et al., 2017). These issues are certainly important. But an open, highly powered, and replicable experiment might nevertheless contain a fatal confound or other source of uncertainty that prevents it from generating novel insights. Therefore, the effort to understand the mechanistic basis of intelligent behavior will continue to involve—as it long has (Boring, 1954)—innovations in the experimental control of phenomena that reside deep within our minds.

REFERENCES

Abrahamyan, A., Silva, L.L., Dakin, S.C., Carandini, M., and Gardner, J.L. (2016). Adaptable history biases in human perceptual decisions. Proc. Natl. Acad. Sci. USA *113*, E3548–E3557.

Akrami, A., Kopec, C.D., Diamond, M.E., and Brody, C.D. (2018). Posterior parietal cortex represents sensory history and mediates its effects on behaviour. Nature *554*, 368–372.

Anticevic, A., Cole, M.W., Murray, J.D., Corlett, P.R., Wang, X.J., and Krystal, J.H. (2012). The role of default network deactivation in cognition and disease. Trends Cogn. Sci. *16*, 584–592.

Bechtel, W. (2007). Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience (Routledge).

Bode, S., and Haynes, J.D. (2009). Decoding sequential stages of task preparation in the human brain. Neuroimage *45*, 606–613.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J.D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. Psychol. Rev. *113*, 700–765.

Boring, E.G. (1954). The nature and history of experimental control. Am. J. Psychol. *67*, 573–589.

Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., and Cohen, J.D. (2001). Conflict monitoring and cognitive control. Psychol. Rev. *108*, 624–652.

Boynton, G.M., Engel, S.A., Glover, G.H., and Heeger, D.J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. J. Neurosci. *16*, 4207–4221.

Braver, T.S., Cohen, J.D., Nystrom, L.E., Jonides, J., Smith, E.E., and Noll, D.C. (1997). A parametric study of prefrontal cortex involvement in human working memory. Neuroimage *5*, 49–62.

Brown, J.W. (2011). Medial prefrontal cortex activity correlates with time-on-task: what does this tell us about theories of cognitive control? Neuroimage *57*, 314–315.

Brunton, B.W., Botvinick, M.M., and Brody, C.D. (2013). Rats and humans can optimally accumulate evidence for decision-making. Science *340*, 95–98.

Buckner, R.L., Andrews-Hanna, J.R., and Schacter, D.L. (2008). The brain's default network: anatomy, function, and relevance to disease. Ann. N Y Acad. Sci. *1124*, 1–38.

Burnham, K.P., and Anderson, D.R. (1998). Practical use of the information-theoretic approach. In Model Selection and Inference (Springer), pp. 75–117.

Busemeyer, J., and Diederich, A. (2010). Cognitive modeling (Sage).

Cabeza, R., and Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. J. Cogn. Neurosci. *12*, 1–47.

Carandini, M. (2012). From circuits to behavior: a bridge too far? Nat. Neurosci. *15*, 507–509.

Cheadle, S., Wyart, V., Tsetsos, K., Myers, N., de Gardelle, V., Herce Castañón, S., and Summerfield, C. (2014). Adaptive gain control during human perceptual choice. Neuron *81*, 1429–1441.

Churchland, A.K., and Kiani, R. (2016). Three challenges for connecting model to mechanism in decision-making. Curr. Opin. Behav. Sci. *11*, 74–80.

Corbetta, M., and Shulman, G.L. (2002). Control of goal-directed and stimulus-driven attention in the brain. Nat. Rev. Neurosci. *3*, 201–215.

Cornsweet, T.N. (1962). The staircrase-method in psychophysics. Am. J. Psychol. *75*, 485–491.

Danielmeier, C., and Ullsperger, M. (2011). Post-error adjustments. Front. Psychol. *2*, 233.

Drugowitsch, J., Moreno-Bote, R., Churchland, A.K., Shadlen, M.N., and Pouget, A. (2012). The cost of accumulating evidence in perceptual decision making. J. Neurosci. *32*, 3612–3628.

Drugowitsch, J., Wyart, V., Devauchelle, A.D., and Koechlin, E. (2016). Computational Precision of Mental Inference as Critical Source of Human Choice Suboptimality. Neuron *92*, 1398–1411.

Farrell, S., and Lewandowsky, S. (2018). Computational modeling of cognition and behavior (Cambridge University Press).

Fechner, G.T., Boring, E.G., and Howes, D.H. (1966). Elements of psychophysics (Holt, Rinehard, and Winston).

Fedorenko, E., Duncan, J., and Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. Proc. Natl. Acad. Sci. USA *110*, 16616–16621.

Fetsch, C.R. (2016). The importance of task design and behavioral control for understanding the neural basis of cognitive functions. Curr. Opin. Neurobiol. *37*, 16–22.

Friston, K.J., Price, C.J., Fletcher, P., Moore, C., Frackowiak, R.S.J., and Dolan, R.J. (1996). The trouble with cognitive subtraction. Neuroimage *4*, 97–104.

Gallivan, J.P., Chapman, C.S., Wolpert, D.M., and Flanagan, J.R. (2018). Decision-making in sensorimotor control. Nat. Rev. Neurosci. *19*, 519–534.

Geisler, W.S. (2011). Contributions of ideal observer theory to vision research. Vision Res. *51*, 771–781.

Ghose, G.M., and Maunsell, J.H.R. (2002). Attentional modulation in visual cortex depends on task timing. Nature *419*, 616–620.

Glimcher, P.W. (2003). Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics (MIT Press).

Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. Annu. Rev. Neurosci. *30*, 535–574.

Gold, J.I., and Stocker, A.A. (2017). Visual Decision-Making in an Uncertain and Dynamic World. Annu. Rev. Vis. Sci. *3*, 227–250.

Green, D.M., and Swets, J.A. (1966). Signal Detection Theory and Psychophysics (Peninsula Publishing).

Griffiths, T.L., Chater, N., Kemp, C., Perfors, A., and Tenenbaum, J.B. (2010). Probabilistic models of cognition: exploring representations and inductive biases. Trends Cogn. Sci. *14*, 357–364.

Grinband, J., Savitskaya, J., Wager, T.D., Teichert, T., Ferrera, V.P., and Hirsch, J. (2011). The dorsal medial frontal cortex is sensitive to time on task, not response conflict or error likelihood. Neuroimage *57*, 303–311.

Hanks, T.D., and Summerfield, C. (2017). Perceptual Decision Making in Rodents, Monkeys, and Humans. Neuron *93*, 15–31.

Hanks, T., Kiani, R., and Shadlen, M.N. (2014). A neural mechanism of speed-accuracy tradeoff in macaque area LIP. eLife *3*, e02260.

Hawkins, G.E., Forstmann, B.U., Wagenmakers, E.J., Ratcliff, R., and Brown, S.D. (2015). Revisiting the evidence for collapsing boundaries and urgency signals in perceptual decision-making. J. Neurosci. *35*, 2476–2484.

Haynes, J.D. (2015). A Primer on Pattern-Based Approaches to fMRI: Principles, Pitfalls, and Perspectives. Neuron *87*, 257–270.

Heitz, R.P., and Schall, J.D. (2012). Neural mechanisms of speed-accuracy tradeoff. Neuron *76*, 616–628.

Hernández, A., Zainos, A., and Romo, R. (2000). Neuronal correlates of sensory discrimination in the somatosensory cortex. Proc. Natl. Acad. Sci. USA *97*, 6191–6196.

Huk, A.C., Katz, L.N., and Yates, J.L. (2017). The Role of the Lateral Intraparietal Area in (the Study of) Decision Making. Annu. Rev. Neurosci. *40*, 349–372.

Jaynes, E.T. (2003). Probability theory: The logic of science (Cambridge University Press).

Kane, M.J., Conway, A.R.A., Miura, T.K., and Colflesh, G.J.H. (2007). Working memory, attention control, and the N-back task: a question of construct validity. J. Exp. Psychol. Learn. Mem. Cogn. *33*, 615–622.

Kepecs, A., Uchida, N., Zariwala, H.A., and Mainen, Z.F. (2008). Neural correlates, computation and behavioural impact of decision confidence. Nature *455*, 227–231.

Khalvati, K., and Rao, R.P. (2015). A Bayesian framework for modeling confidence in perceptual decision making. In NIPS'15 Proceedings of the 28th International Conference on Neural Information Processing Systems 2, pp. 2413–2421.

Kiani, R., and Shadlen, M.N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. Science *324*, 759–764.

Kiani, R., Hanks, T.D., and Shadlen, M.N. (2008). Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. J. Neurosci. *28*, 3017–3029.

Kingdom, F.A., and Prins, N. (2010). Psychophysics: A Practical Introduction (Elsevier).

Kolling, N., Behrens, T.E.J., Mars, R.B., and Rushworth, M.F.S. (2012). Neural mechanisms of foraging. Science *336*, 95–98.

Kolling, N., Behrens, T., Wittmann, M.K., and Rushworth, M. (2016). Multiple signals in anterior cingulate cortex. Curr. Opin. Neurobiol. *37*, 36–43.

Krakauer, J.W., Ghazanfar, A.A., Gomez-Marin, A., MacIver, M.A., and Poeppel, D. (2017). Neuroscience Needs Behavior: Correcting a Reductionist Bias. Neuron *93*, 480–490.

Link, S.W. (1992). The Wave Theory of Difference and Similarity (Lawrence Erlbaum).

Logothetis, N.K., and Wandell, B.A. (2004). Interpreting the BOLD signal. Annu. Rev. Physiol. *66*, 735–769.

Ma, W.J., Husain, M., and Bays, P.M. (2014). Changing concepts of working memory. Nat. Neurosci. *17*, 347–356.

Marmarelis, P.Z., and Naka, K. (1972). White-noise analysis of a neuron chain: an application of the Wiener theory. Science *175*, 1276–1278.

Marr, D.C. (1982). Vision (MIT Press).

Mars, R.B., Shea, N.J., Kolling, N., and Rushworth, M.F.S. (2012). Model-based analyses: Promises, pitfalls, and example applications to the study of cognitive control. Q J Exp Psychol (Hove) *65*, 252–267.

McDougle, S.D., Ivry, R.B., and Taylor, J.A. (2016). Taking Aim at the Cognitive Side of Learning in Sensorimotor Adaptation Tasks. Trends Cogn. Sci. *20*, 535–544.

Miller, E.K., and Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. Annu. Rev. Neurosci. *24*, 167–202.

Miller, P., and Katz, D.B. (2013). Accuracy and response-time distributions for decision-making: linear perfect integrators versus nonlinear attractor-based neural circuits. J. Comput. Neurosci. *35*, 261–294.

Moran, R. (2015). Optimal decision making in heterogeneous and biased environments. Psychon. Bull. Rev. *22*, 38–53.

Munafò, M.R., Nosek, B.A., Bishop, D.V.M., Button, K.S., Chambers, C.D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.J., Ware, J.J., and Ioannidis, J.P.A. (2017). A manifesto for reproducible science. Nat. Hum. Behav. *1*, e124.

Najafi, F., and Churchland, A.K. (2018). Perceptual Decision-Making: A Field in the Midst of a Transformation. Neuron *100*, 453–462.

Neri, P., and Heeger, D.J. (2002). Spatiotemporal mechanisms for detecting and identifying image features in human vision. Nat. Neurosci. *5*, 812–816.

Nienborg, H., and Cumming, B.G. (2009). Decision-related activity in sensory neurons reflects more than a neuron's causal effect. Nature *459*, 89–92.

Nobre, A.C., and van Ede, F. (2018). Anticipated moments: temporal structure in attention. Nat. Rev. Neurosci. *19*, 34–48.

Okazawa, G., Sha, L., Purcell, B.A., and Kiani, R. (2018). Psychophysical reverse correlation reflects both sensory and decision-making processes. Nat. Commun. *9*, 3479.

Palmer, J., Huk, A.C., and Shadlen, M.N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. J. Vis. *5*, 376–404.

Parker, A.J., and Newsome, W.T. (1998). Sense and the single neuron: probing the physiology of perception. Annu. Rev. Neurosci. *21*, 227–277.

Pisupati, S., Chartarifsky-Lynn, L., Khanal, A., and Churchland, A.K. (2019). Lapses in perceptual decisions reflect exploration. bioRxiv. https://doi.org/10.1101/613828.

Poldrack, R.A. (2011). Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. Neuron *72*, 692–697.

Purcell, B.A., and Kiani, R. (2016). Neural Mechanisms of Post-error Adjustments of Decision Policy in Parietal Cortex. Neuron *89*, 658–671.

Quick, R.F., Jr. (1974). A vector-magnitude model of contrast detection. Kybernetik *16*, 65–67.

Rabbitt, P., and Rodgers, B. (1977). What does a man do after he makes an error? An analysis of response programming. Q. J. Exp. Psychol. *29*, 727–743.

Rahnev, D., and Denison, R.N. (2018). Suboptimality in Perceptual Decision Making. Behav. Brain Sci. *41*, 1–107.

Raichle, M.E. (2015). The brain's default mode network. Annu. Rev. Neurosci. *38*, 433–447.

Raposo, D., Kaufman, M.T., and Churchland, A.K. (2014). A category-free neural population supports evolving demands during decision-making. Nat. Neurosci. *17*, 1784–1792.

Ratcliff, R., and Rouder, J.N. (1998). Modeling Response Times for Two-Choice Decisions. Psychol. Sci. *9*, 347–356.

Ratcliff, R., Smith, P.L., Brown, S.D., and McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. Trends Cogn. Sci. *20*, 260–281.

Roitman, J.D., and Shadlen, M.N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. J. Neurosci. *22*, 9475–9489.

Rust, N.C., and Movshon, J.A. (2005). In praise of artifice. Nat. Neurosci. *8*, 1647–1650.

Schall, J.D. (2001). Neural basis of deciding, choosing and acting. Nat. Rev. Neurosci. *2*, 33–42.

Scott, B.B., Constantinople, C.M., Akrami, A., Hanks, T.D., Brody, C.D., and Tank, D.W. (2017). Fronto-parietal Cortical Circuits Encode Accumulated Evidence with a Diversity of Timescales. Neuron *95*, 385–398.

Servant, M., Tillman, G., Schall, J.D., Logan, G.D., and Palmeri, T.J. (2019). Neurally constrained modeling of speed-accuracy tradeoff during visual search: gated accumulation of modulated evidence. J. Neurophysiol. *121*, 1300–1314.

Shadlen, M.N., and Kiani, R. (2013). Decision making as a window on cognition. Neuron *80*, 791–806.

Shadlen, M.N., and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J. Neurophysiol. *86*, 1916–1936.

Shadlen, M.N., Hanks, T.D., Churchland, A.K., Kiani, R., and Yang, T. (2006). The Speed and Accuracy of a Simple Perceptual Decision: A Mathematical Primer. In Bayesian Brain: Probabilistic Approaches to Neural Coding, K. Doya, ed. (MIT Press).

Shenhav, A., Straccia, M.A., Cohen, J.D., and Botvinick, M.M. (2014). Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. Nat. Neurosci. *17*, 1249–1254.

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T.L., Cohen, J.D., and Botvinick, M.M. (2017). Toward a Rational and Mechanistic Account of Mental Effort. Annu. Rev. Neurosci. *40*, 99–124.

Smith, P.L., and Little, D.R. (2018). Small is beautiful: In defense of the small-N design. Psychon. Bull. Rev. *25*, 2083–2101.

Stocker, A.A. (2018). Credo for optimality. Behav. Brain Sci. *41*, e244.

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.

Summerfield, C., and Tsetsos, K. (2015). Do humans make good decisions? Trends Cogn. Sci. *19*, 27–34.

Teller, D.Y. (1984). Linking propositions. Vision Res. *24*, 1233–1246.

Thura, D., and Cisek, P. (2016). On the difference between evidence accumulator models and the urgency gating model. J. Neurophysiol. *115*, 622–623.

Todd, M.T., Nystrom, L.E., and Cohen, J.D. (2013). Confounds in multivariate pattern analysis: Theory and rule representation case study. Neuroimage *77*, 157–165.

Tsetsos, K., Gao, J., McClelland, J.L., and Usher, M. (2012). Using Time-Varying Evidence to Test Models of Decision Dynamics: Bounded Diffusion vs. the Leaky Competing Accumulator Model. Front. Neurosci. *6*, 79.

Turner, B.M., Miletić, S., and Forstmann, B.U. (2018). Outlook on deep neural networks in computational cognitive neuroscience. Neuroimage *180* (Pt A), 117–118.

Tversky, A., and Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. Science *185*, 1124–1131.

van Ravenzwaaij, D., Provost, A., and Brown, S. (2017). A confirmatory approach for integrating neural and behavioral data into a single model. J. Math. Psychol. *76*, 131–141.

von Helmholtz, H. (1867). Handbuch der physiologischen optik (Leopold Voss).

Wang, X.J., and Krystal, J.H. (2014). Computational psychiatry. Neuron *84*, 638–654.

Waskom, M.L., and Kiani, R. (2018). Decision Making through Integration of Sensory Evidence at Prolonged Timescales. Curr. Biol. *28*, 3850–3856.

Waskom, M.L., and Wagner, A.D. (2017). Distributed representation of context by intrinsic subnetworks in prefrontal cortex. Proc. Natl. Acad. Sci. USA *114*, 2030–2035.

Wolpert, D.M., Diedrichsen, J., and Flanagan, J.R. (2011). Principles of sensorimotor learning. Nat. Rev. Neurosci. *12*, 739–751.

Woolgar, A., Hampshire, A., Thompson, R., and Duncan, J. (2011). Adaptive coding of task-relevant information in human frontoparietal cortex. J. Neurosci. *31*, 14592–14599.

Woolgar, A., Golland, P., and Bode, S. (2014). Coping with confounds in multivoxel pattern analysis: what should we do about reaction time differences? A comment on Todd, Nystrom & Cohen 2013. Neuroimage *98*, 506–512.

Yarkoni, T., Poldrack, R.A., Nichols, T.E., Van Essen, D.C., and Wager, T.D. (2011). Large-scale automated synthesis of human functional neuroimaging data. Nat. Methods *8*, 665–670.

Yeung, N., Cohen, J.D., and Botvinick, M.M. (2011). Errors of interpretation and modeling: a reply to Grinband et al. Neuroimage *57*, 316–319.