Published in final edited form as:

J Math Psychol. 2017 February ; 76(Pt B): 59–64. doi:10.1016/j.jmp.2016.10.010.

Model-based cognitive neuroscience

Thomas J. Palmeri[a,*], **Bradley C. Love**[b], and **Brandon M. Turner**[c]

[a]Vanderbilt University, United States

[b]University College London, United Kingdom

[c]The Ohio State University, United States

Abstract

This special issue explores the growing intersection between mathematical psychology and cognitive neuroscience. Mathematical psychology, and cognitive modeling more generally, has a rich history of formalizing and testing hypotheses about cognitive mechanisms within a mathematical and computational language, making exquisite predictions of how people perceive, learn, remember, and decide. Cognitive neuroscience aims to identify neural mechanisms associated with key aspects of cognition using techniques like neurophysiology, electrophysiology, and structural and functional brain imaging. These two come together in a powerful new approach called *model-based cognitive neuroscience*, which can both inform cognitive modeling and help to interpret neural measures. Cognitive models decompose complex behavior into representations and processes and these latent model states can be used to explain the modulation of brain states under different experimental conditions. Reciprocally, neural measures provide data that help constrain cognitive models and adjudicate between competing cognitive models that make similar predictions about behavior. As examples, brain measures are related to cognitive model parameters fitted to individual participant data, measures of brain dynamics are related to measures of model dynamics, model parameters are constrained by neural measures, model parameters or model states are used in statistical analyses of neural data, or neural and behavioral data are analyzed jointly within a hierarchical modeling framework. We provide an introduction to the field of model-based cognitive neuroscience and to the articles contained within this special issue.

Keywords

Cognitive modeling; Cognitive neuroscience; Model-based cognitive neuroscience

Exciting new synergies between mathematical psychology and cognitive neuroscience have emerged. This special issue of the *Journal of Mathematical Psychology* includes reviews, tutorials, and original research papers highlighting this new area of *model-based cognitive neuroscience*. In this opening article, we outline this new approach and introduce the articles contained in this special issue.

[*]Correspondence to: Department of Psychology, 301 Wilson Hall, Vanderbilt University, Nashville, TN 37240, United States. thomas.j.palmeri@vanderbilt.edu (T.J. Palmeri).

1. What is model-based cognitive neuroscience?

Alternative approaches to theory in both psychology and neuroscience often begin by considering Marr's (1982) classic three levels: The *computational level* considers the goals of the organism and the structure of the environment, without considering mechanism, typified by many Bayesian theories of the mind (e.g., Anderson, 1990; Oaksford & Chater, 2007; Tenenbaum, Kemp, Griffiths, & Goodman, 2011). The *algorithmic level* considers what representations and processes underlie cognition and perception, without considering their biological realization, typified by many mathematical and computational models of cognition and perception (e.g., Busemeyer, Wang, Townsend, & Eidels, 2015; Sun, 2008). The *implementation level* asks how mechanisms are physically realized within a biological substrate, namely neurons and their connections in the brain, typified by classical theoretical work in neuroscience (e.g., Carnevale & Hines, 2006; Dayan & Abbott, 2005).

While Marr envisioned connections between these levels, there often had been intellectual and disciplinary barriers to considering explanations that crossed levels. This led theorists to work traditionally within only one level of analysis. Not so long ago, a graduate student trained in mathematical psychology considering postdoctoral training in neuroscience might have been about as sensible as considering running off to join the circus. For some, the brain could well be made of tinker toys for its relevance to understanding human cognition. As well, not so long ago, few trained in systems neuroscience would ever consider whether insights from cognitive and mathematical psychology might inform understanding of neural function. Cognitive conceptual building blocks were often thought little more than folk psychology, with philosophical arguments lending support to a strict reductionist approach to understanding the brain (e.g., Churchland, 1986).[1]

Early attempts to address this impasse focused on connectionist models of cognition that took inspiration from the brain. Connectionists viewed the brain as consisting of simple computing units (akin to neurons) that integrated signals passed across connection weights that were adjusted by learning rules. However, these models rarely made contact with the implementational details of the brain. In most cases these models served as existence proofs that a model consisting of many simple computing elements could accomplish a task in roughly the same fashion as a human. Nevertheless, these models were attempts to bridge levels of analysis and were championed as more biologically plausible than competing models at the algorithmic or computational levels. Unfortunately, notions of biological plausibility were rarely defined nor evaluated rigorously. The gap between levels of analyses stubbornly remained.

Model-based cognitive neuroscience breaks the traditional barriers between models and the brain (e.g., Forstmann, Wagenmakers, Eichele, Brown, & Serences, 2011; Forstmann, 2015; Palmeri, Schall, Logan, & Townsend, 2015; Smith & Ratcliff, 2004). From the perspective

[1]Of course, there were exceptions to barriers between the algorithmic level and the implementation level, to again cast this in Marr's terms. In the case of relatively low-level visual sensation and perception, there have long been deep connections between theoretical work in visual psychophysics and the underlying visual neurophysiology and neuroanatomy, in part because the relevant neural hardware is not far removed from the source of visual stimulation. And the field of cognitive neuropsychology has long considered theoretically how cases of brain damage and neurodegenerative and neurodevelopmental disorders influence understanding of human cognition.

of cognitive and mathematical psychology, formal models explain behavior in terms of representations and processes instantiated in mathematics and computations, and observed variation in behavior across experimental conditions and individuals is explained in terms of variation in model parameters and model states. Model-based cognitive neuroscience allows for consideration of whether these latent model parameters or model states might be related to, or constrained by, observed brain measures or brain states, over and above whether a model fits or predicts observed behavior. From the perspective of systems and cognitive neuroscience, a key component of understanding neurons, neural circuits, or brain areas is explaining the computations that they perform. In a model-based cognitive neuroscience approach, to the extent that brain measures or brain states are predicted by model parameters or model states, those models provide a potential explanation of brain function, regardless of whether or not those models are implemented in neuron-like elements.

The emergence and growth of model-based cognitive neuroscience over the past decade can be attributed to a number of converging forces. One was the recognition on the part of cognitive modelers and mathematical psychologists interested in understanding the mechanisms that brain data is simply additional data by which to constrain and contrast models. Response probabilities, response times, confidence ratings and the like are the outcomes of processing. Brain data reflect intermediary states. Considering how internal processes predicted by a model relate to internal processes measured in the brain can break theoretical stalemates caused by model mimicry. While two different models making different mechanistic assumptions about representations and processes may make similar predictions about observed behavior, they may well make different predictions about internal model states, which can then be compared with or constrained by measured brain states (e.g., Boucher, Palmeri, Logan, & Schall, 2007; Mack, Preston, & Love, 2013; Palmeri, 2014; Purcell et al., 2010; Purcell, Schall, Logan, & Palmeri, 2012).

Another force was the recognition on the part of cognitive and systems neuroscientists for the need for new approaches to making sense of the growing body of neural data from functional brain imaging, electrophysiology, neurophysiology, and other neuroscience techniques. Correlating brain measures with stimuli, conditions, and responses provides only a rather limited window on understanding brain function. To go beyond merely mapping out which brain areas or which neurons modulate their activity under which conditions means to explain and understand what mechanisms and computations are engaged within those brain areas or neurons. Algorithmic and computational models provide a language and a body of viable hypotheses, as well as a set of tools, for explaining and understanding those neural mechanisms and computations.

Recognition has grown for considering the algorithms and computations that underlie neural processing. Carandini (2012) characterized any direct link between neural circuits and behavior as a “bridge too far”, and argued that it was necessary to theorize at an intermediate level in Marr’s hierarchy, considering the algorithms and computations that neural circuits perform. The purely bottom-up approach to understanding the brain that characterized the initial stages of the billion Euro Human Brain Project was widely criticized by cognitive and computational neuroscientists and led to a shake-up of its leadership and vision (e.g., Enserink & Kupferschmidt, 2014; Theil, 2015). Rather than adopting a strictly bottom-up

(or top-down) approach, model-based cognitive neuroscience can be characterized as an inside-out approach (Love, 2015), that may well be a level of theorizing that is just right (Logan, Yamaguchi, Schall, & Palmeri, 2015).

Perhaps the most potent force propelling model-based cognitive neuroscience over the past decade has been its demonstrated success in providing new insight at both the cognitive and neural levels. One especially salient body of work has centered around accumulator models of decision making, a well-known class of models with a long history in cognitive psychology (e.g., Ratcliff & Smith, 2004). These models assume that variability in choice probability and response times arise from variability in the, often noisy, accumulation of evidence to response thresholds, and variants of these models have accounted for decisions in perception, memory, categorization, and other tasks (e.g., Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Brown & Heathcote, 2008; Forstmann, Ratcliff, & Wagenmakers, 2016; Nosofsky & Palmeri, 1997; Palmeri, 1997). As one of the first examples of systems neuroscience making contact with cognitive modeling, when Hanes and Schall (1996) were interested in understanding how neurons in Frontal Eye Field (FEF) decide where and when to saccade in the visual field, they turned to the cognitive modeling literature for inspiration and insight. Based on the fact that the dynamics of certain FEF neurons mirrored the dynamics in accumulators, accumulation of evidence models provided a language for describing the computations that these FEF neurons were engaged in. Cognitive models provided insight into neural processes. Hanes and Schall also showed that the dynamics of these FEF neurons were more consistent with variable accumulation to a fixed threshold than fixed accumulation to a variable threshold, two competing mechanistic hypotheses of decision making that can be difficult to distinguish based on behavioral data alone (Grice, 1968). Neural data provided insight into cognitive models.

These initial insights spawned a considerable body of research linking neurophysiology and cognitive modeling to understand elementary decision making (e.g., Forstmann et al., 2016; Gold & Shadlen, 2007; Logan et al., 2015; Mazurek, Roitman, Ditterich, & Shadlen, 2003; Palmeri et al., 2015; Ratcliff, Cherian, & Segraves, 2003; Schall, 2001, 2004; Smith & Ratcliff, 2004; Zandbelt, Purcell, Palmeri, Logan, & Schall, 2014). As one illustrative example, Purcell et al. (2010, 2012) applied accumulator models to understand existing data on the behavior and neurophysiology of saccade decision making by awake behaving primates (e.g., Bichot & Schall, 1999; Cohen, Heitz, Woodman, & Schall, 2009). Adopting a classic approach used in mathematical psychology, they formulated a variety of alternative models assuming various architectural components characteristic of various accumulator models of decision making, rejecting models that could not account qualitatively and quantitatively for observed response probabilities and distribution of response times for saccades.

Going beyond a pure mathematical psychology approach of fitting models to behavioral data, they turned to neurophysiology in two ways. First, they allowed neurophysiology to constrain key model components. In many, but not all (e.g., Nosofsky & Palmeri, 1997; Palmeri, 1997), applications of accumulator models, the rate at which evidence is accumulated, the *drift rate*, is allowed to be a free parameter. Purcell et al. (2010, 2012) instead instantiated an hypothesis that a particular class of neurons in FEF (*visually-*

responsive neurons) represent the evidence that is accumulated, replacing the drift rate and other parameters with recorded neurophysiology. Neurophysiology significantly limited the flexibility of various model architectures to account for observed behavioral data.

Second, faced with several alternative model architectures that could account equally well for the observed behavioral data, if they had no other data to turn to, they would have had to appeal to parsimony in selecting a winning model architecture (see also Boucher et al., 2007; Logan et al., 2015). Purcell et al. (2010, 2012) instead turned to neurophysiology as an additional data source for contrasting between alternative models. Adopting the linking proposition (Schall, 2004; Teller, 1984) that *movement-related neurons* in FEF instantiate an accumulation of evidence to a threshold for saccade decisions (Hanes & Schall, 1996), they compared the predicted dynamics of model accumulation to the observed dynamics of these FEF neurons (see also Purcell & Palmeri, 2017). Only their gated accumulator model could both account for the behavioral data and predict the dynamics of FEF movement-related neurons. Neurophysiology provided key data by which to contrast models that otherwise provided the same predictions of overt behavior.

Another approach for avoiding the theoretical stalemate that can ensue when fitting complex models to behavioral data alone is to treat neural data as auxiliary information on which latent model mechanisms should covary. Models like the classic diffusion model (Ratcliff & Rouder, 1998; Ratcliff & Smith, 2004) have three sources of trial-to-trial variability, assuming fluctuations in things like response bias, the rate of evidence accumulation, and perceptual and motor non-decision time. The assumption is that these parameters vary from one trial to another in ways that are completely consistent throughout the experiment (an assumption known as independent, identically distributed). However, because there is no mechanism to guide these fluctuations, we cannot appreciate aspects of the decision process that are vital to ensuring success on a specific trial. Furthermore, these assumptions are at odds with several findings in neuroscience that implicate the gradual waxing and waning of attention on behavioral performance. In summary of these findings, unique networks of brain activity arise from separating neural data on the basis of behavioral data: an “off-task” network gives rise to poor behavioral performance whereas an “on-task” network gives rise to good behavioral performance (Mittner et al., 2014; Turner, Van Maanen, & Forstmann, 2015).

These observations lead Turner et al. (2015) to develop a model that blends neuroscience and mathematical psychology to formally ground decision-making models with neurophysiology. Their strategy was to treat trial-by-trial neural data (as measured by fMRI) as information about the trial-to-trial fluctuations in the latent parameters assumed by a standard diffusion model. The model was constructed on the basis of a previously developed framework for imposing neurophysiological constraints on behavioral models across subjects (Turner, 2015; Turner et al., 2013), but extends this framework to a trial-by-trial basis. Once fit to data, the model was able to articulate how disparate networks of brain activity were associated with orthogonal mechanisms in the model, such as pre-stimulus bias and the rate of evidence accumulation. Turner et al. also showed that not only did their model provide a new perspective on both neural and behavioral data using generative

modeling techniques, but the model could also outperform a standard diffusion model that only considered behavioral data in a leave-one-out cross-validation test.

Model-based neuroscience opens up possibilities for cognitive models to take on second lives as formal neuroscientific theories. For example, Love and Gureckis (2007) proposed a theory linking aspects of the SUSTAIN clustering model of human categorization (Love, Medin, & Gureckis, 2004; Sakamoto & Love, 2004) to the functions of prefrontal cortex and the hippocampus. They simulating various populations, such as amnesics (Love & Gureckis, 2007), infants (Gureckis & Love, 2004), and the aged (Davis, Love, & Maddox, 2012), by adjusting model parameters hypothesized to relate to brain regions whose functions vary across populations. With the advent of model-based neuroscience, exact predictions of the theory were tested and confirmed with healthy young adults using fMRI (Davis et al., 2012; Davis, Love, & Preston, 2012a; Davis, Xue, Love, Preston, & Poldrack, 2014; Mack, Preston, & Love, in press). The analyses revealed a number of phenomena that would not be possible to observe without the model, such as how the involvement of the hippocampus changes over learning trials, ramping up for familiar items (related to recognition) at the time of decision and ramping down at the time of feedback as the error signal abates (Davis et al., 2012). The model-based imaging work also confirmed more speculative hypotheses such as that prefrontal and hippocampus interactions would be strongest in the early stages of mastering a new learning task as attention weights are established (Mack et al., in press).

While we opened this article by contrasting a bottom-up neural network approach with an inside-out (Love, 2015) cognitive modeling approach to relating brain and behavior, we want to make clear that the approaches used in a model-based cognitive neuroscience can just as well be applied to neural network models as to more abstract cognitive models. The SUSTAIN model (Love et al., 2004) used by Davis et al. (2012,a); Davis, Love, and Preston (2012b) discussed above is instantiated using a number of neural network building blocks. Yet the relation between SUSTAIN and brain imaging data is not cemented by any mapping from neural-like model elements to neurons in the brain, but by the ability of patterns of activity in the model to reveal and explain patterns of activity in the brain. Similarly, the overall structure of so-called deep learning models of vision (e.g., LeCun, Bengio, & Hinton, 2015) are inspired by neural networks and key aspects of the neurophysiology of the primate visual system. But the insights provided by these models into understanding the representation of objects in the brain (Kriegeskorte, 2015; Yamins et al., 2014) is based on how well patterns within high-level representational layers of these models predict patterns of brain activity, not on the neural-like building blocks of these models (Khaligh-Razavi, Henriksson, Kay, & Kriegeskorte, 2017; Khaligh-Razavi & Kriegeskorte, 2014).

2. Overview

Here we provide brief outlines of the papers that appear in this special issue:

Approaches to analysis in model-based cognitive neuroscience

Turner, Forstmann, Love, Palmeri, and Van Maanen (2017) provide an overarching framework for describing the varying approaches to model-based cognitive neuroscience that have emerged in the literature over the past several years. They organize these

approaches on the basis of particular theoretical goals, which include using neural data to constrain a cognitive model, using a cognitive model to predict neural data, and accounting for both neural and behavioral data simultaneously using the same model. Accompanying each of these theoretical goals, they highlight some particularly successful examples. They also provide a conceptual guide to choosing among various approaches when performing model-based cognitive neuroscience.

Integrating theoretical models with functional neuroimaging

Pratte and Tong (2017) highlight a number of salient examples linking cognitive models and functional brain imaging data using a model-based cognitive neuroscience approach. Their selective review spans a broad range of core topics in perception and cognition, including visual perception (Brouwer & Heeger, 2011), attention (Pratte, Ling, Swisher, & Tong, 2013), long-term memory (Kragel, Morton, & Polyn, 2015), categorization (Mack et al., 2013), and cognitive control (Ide, Shenoy, Yu, & Li, 2013).

A step-by-step tutorial on using the cognitive architecture ACT-R in combination with fMRI data

Borst and Anderson (2017) provide a tutorial on using the ACT-R cognitive architecture (e.g., Anderson, Bothell, Lebiere, & Matessa, 1998) to understand fMRI data (e.g., Anderson, Betts, Ferris, & Fincham, 2010; Anderson, Fincham, Qin, & Stocco, 2008; Borst & Anderson, 2013). They illustrate how ACT-R can be used in combination with fMRI data in two different ways: first that fMRI data can be used to evaluate and constrain models in ACT-R by means of predefined Region-of-Interest (ROI) analysis, and second that predictions from ACT-R models can be used to locate neural correlates of model processes and representations by means of model-based fMRI analysis. As a tutorial, they provide code and worked examples of both types of analysis on a math problem solving task performed in an fMRI scanner.

Variability in behavior that cognitive models do not explain can be linked to neuroimaging data

Gluth and Rieskamp (2017) review evidence for the proposal that neural and behavioral variability can be linked to one another by allowing moment-to-moment fluctuations in neural measures, like fMRI and EEG, to inform trial-by-trial variability in cognitive model parameters. One approach to linking single-trial measures of the brain to single-trial parameters in models has been to simply regress them onto one another. Gluth and Rieskamp provide a tutorial of a novel and efficient alternative approach that goes beyond a raw two-stage correlational approach by increasing the resolution of the single-trial parameter estimates in an iterative fashion, similar in some ways to an EM algorithm. As illustration, they show how the variability in the parameters of an accumulator (sequential sampling) model can be related to variability in neuroimaging data.

How attention influences perceptual decision making: single-trial EEG correlates of drift-diffusion model parameters

Nunez, Vandekerckhove, and Srinivasan (2017) provide a specific illustration of how variability in neural measures can constrain variability in model parameters. Within a hierarchical Bayesian framework, various forms of a drift-diffusion model are fitted to behavioral data from a perceptual decision making task, with different model forms assuming different mathematical relationships between model parameters and EEG measures. Trial-to-trial measures of certain key attention-related evoked potentials in simultaneous EEG recordings can explain trial-to-trial evidence accumulation rates and perceptual processing times in a diffusion model fitted to perceptual decision making behavior.

A confirmatory approach for integrating neural and behavioral data into a single model

van Ravenzwaaij, Provost, and Brown (2017) provide another illustration of a joint modeling approach to model-based cognitive neuroscience. Within a hierarchical Bayesian framework they use the Linear Ballistic Accumulator (LBA) model (Brown & Heathcote, 2008) to account for behavioral data during a mental rotation task, testing different hypotheses linking cognitive model parameters and neural data measured via event-related potentials (ERPs). They specifically investigate how changes in drift rate and non-decision time with mental rotation angle might be constrained by changes in certain ERP amplitudes measured during the task.

On the efficiency of neurally-informed cognitive models to identify latent cognitive states

Hawkins, Mittner, Forstmann, and Heathcote (2017) illustrate how neural data can be used to test between cognitive models with different latent states. They focus on whether the underlying states driving performance in a speeded decision tasks are discrete or continuous. Through model recovery studies the authors determine that discrete state models are more robustly recovered than continuous state models, suggesting that neural data may more easily be linked to certain varieties of cognitive models.

Relating accumulator model parameters and neural dynamics

Purcell and Palmeri (2017) build on the work cited earlier on the identification of neural activity in certain brain areas with evidence accumulation in sequential sampling models. Through simulations, they caution against simply equating variability in measures of neural dynamics with variability in cognitive model parameters. Simulated variation in model dynamics in accumulators is not always related one-to-one with variation of accumulator model parameters. The most general mapping between neural measures and model mechanisms may be one between measured neural dynamics and predicted model dynamics, not one between measured neural dynamics and model parameters.

A primer on encoding models in sensory neuroscience

van Gerven (2017) explores fundamental questions of how the primate visual system represents the visual world. In visual neuroscience, the concept of the receptive field has been a key concept for understanding the response properties of neurons. While classical

receptive field mapping has proved successful for understanding representations in early visual areas like area V1, more general methods are needed for understanding higher-level visual representations and to allow for non-invasive mapping of visual representations in humans. van Gerven provides a mathematical and computational primer on Encoding Models, which at first approximation can be described as generalization of classical receptive field and population receptive field approaches, allowing for the nonlinear response properties of complex representations in high-level visual areas and their manifestation in functional brain imaging to be well characterized.

Fixed versus mixed RSA: Explaining visual representations by fixed and mixed feature sets from shallow and deep computational models

Khaligh-Razavi et al. (2017) provide a complementary approach to understanding how the primate visual system represents the world. Their starting point is existing neural network and computer vision models of object recognition. Their question is whether the representations produced in the model predict the activity measured in the brain. Using a technique called Representational Similarity Analysis (Kriegeskorte, 2009, 2015), they ask whether patterns of similarities in object representations produced in particular layers of a model are analogous to patterns of similarities measured in particular areas of the brain. Deep learning models (LeCun et al., 2015) have provided good accounts of object representations observed in high-level visual areas (Khaligh-Razavi & Kriegeskorte, 2014; Kriegeskorte, 2015; Yamins et al., 2014); this article reviews that work and outlines approaches to fixing or mixing the model representations when comparing to brain measures.

A tutorial on the free-energy framework for modeling perception and learning

Bogacz (2017) provides a tutorial on free energy and related predictive coding approaches. In these approaches, models assume that the sensory cortex infers the most likely values of attributes or features of sensory stimuli from the noisy inputs encoding the stimuli. The author demonstrates how powerful inferences can be made by very simple computations that could be carried out by neurons. Clear examples help the reader grasp these general concepts that link measures of uncertainty with neural computations.

Model-based functional neuroimaging using dynamic neural fields: An integrative cognitive neuroscience approach

Wijeakumar, Ambrose, Spencer, and Curtu (2017) provide a review and tutorial of an approach to model-based cognitive neuroscience using a theoretical framework called Dynamic Field Theory (Erlhagen & Schöner, 2002) applied to functional brain imaging (Buss, Wifall, Hazeltine, & Spencer, 2009). They outline the assumptions of DFT and how it is applied to behavioral data, describe how parameters of the model can be used in brain imaging analyses, and compare the model-based cognitive neuroscience approach to standard brain imaging analyses of the same dataset.

3. Guest consulting editors

We thank the ad hoc reviewers who contributed critical comments that helped shape the papers appearing in this special issue; we especially thank the following for serving as guest consulting editors for this special issue: Jelmer Borst (University of Groningen), Tyler Davis (Texas Tech University), Birte Forstmann (University of Amsterdam), Scott Brown (University of Newcastle), Sam Gershman (Harvard University), Laurence Hunt (University College London), Xiaosi Gu (University of Texas at Dallas), Michael Mack (University of Toronto), Neal Morton (University of Texas at Austin), Braden Purcell (New York University), Michael Pratte (Mississippi State University), Per Sederberg (The Ohio State University), Mark Steyvers (University of California Irvine), Marcel van Gerven (Donders Institute), Marieke van Vugt (University of Groningen), Joachim Vandekerckhove (University of California Irvine), Corey White (Syracuse University), Bram Zandbelt (Donders Institute).

Acknowledgments

TJP was supported by NIH R01-EY021833, NSF Temporal Dynamics of Learning Center SMA-1041755, and NIH P30-EY08126; BCL was supported by Leverhulme Trust grant RPG-2014-075, NIH 1P01-HD080679, and a Wellcome Trust Investigator Award WT106931MA.

References

Anderson JR. The adaptive character of thought. Psychology Press; 1990.

Anderson JR, Betts S, Ferris JL, Fincham JM. Neural imaging to track mental states. Proceedings of the National Academy of Sciences of the United States. 2010; 107:7018–7023.

Anderson JR, Bothell D, Lebiere C, Matessa M. An integrated theory of list memory. Journal of Memory and Language. 1998; 38:341–380.

Anderson JR, Fincham JM, Qin Y, Stocco A. A central circuit of the mind. Trends in Cognitive Science. 2008; 12:136–143.

Bichot NP, Schall JD. Effects of similarity and history on neural mechanisms of visual selection. Nature Neuroscience. 1999; 2:549–554. [PubMed: 10448220]

Bogacz R. A tutorial on the free-energy framework for modelling perception and learning. Journal of Mathematical Psychology. 2017; 76:198–211. [PubMed: 28298703]

Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD. The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. Psychological Review. 2006; 113:700–765. [PubMed: 17014301]

Borst JP, Anderson JR. Using model-based functional MRI to locate working memory updates and declarative memory retrievals in the fronto-parietal network. Proceedings of the National Academy of Sciences of the United States. 2013; 110:1628–1633.

Borst JP, Anderson JR. A step-by-step tutorial on using the cognitive architecture ACT-R in combination with fMRI Data. Journal of Mathematical Psychology. 2017; 76:94–103.

Boucher L, Palmeri TJ, Logan GD, Schall JD. Inhibitory control in mind and brain: An interactive race model of countermanding saccades. Psychological Review. 2007; 114:376–397. [PubMed: 17500631]

Brouwer GJ, Heeger DJ. Cross-orientation suppression in human visual cortex. Journal of Neurophysiology. 2011; 106(5):2108–2119. [PubMed: 21775720]

Brown S, Heathcote A. The simplest complete model of choice re-action time: Linear ballistic accumulation. Cognitive Psychology. 2008; 57:153–178. [PubMed: 18243170]

Busemeyer JR, Wang Z, Townsend JT, Eidels A. The Oxford handbook of computational and mathematical psychology. Oxford University Press; 2015.

Buss AT, Wifall T, Hazeltine E, Spencer JP. Integrating the behavioral and neural dynamics of response selection in a dual-task paradigm: A dynamic neural field model of Dux et al. (2009). Journal of Cognitive Neuroscience. 2009; 26(2):334–351.

Carandini M. From circuits to behavior: a bridge too far? Nature Neuroscience. 2012; 15:507–509. [PubMed: 22449960]

Carnevale NT, Hines ML. The NEURON book. Cambridge, UK: Cambridge University Press; 2006.

Churchland PS. Neurophilosophy: toward a unified science of the mind/brain. Cambridge, MA: MIT Press; 1986.

Cohen JY, Heitz RP, Woodman GF, Schall JD. Neural basis of the set-size effect in frontal eye field: Timing of attention during visual search. Journal of Neurophysiology. 2009; 101:1699–1704. [PubMed: 19176607]

Davis T, Love BC, Maddox WT. Age-related declines in the fidelity of newly acquired category representations. Learning and Memory. 2012; 19:325–329. [PubMed: 22815536]

Davis T, Love BC, Preston AR. Learning the exception to the rule: Model-based fMRI reveals specialized representations for surprising category members. Cerebral Cortex. 2012a; 22:260–273. [PubMed: 21666132]

Davis T, Love BC, Preston AR. Striatal and hippocampal entropy and recognition signals in category learning: Simultaneous processes revealed by model-based fMRI. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2012b; 38:821–839.

Davis T, Xue G, Love BC, Preston AR, Poldrack RA. Global neural pattern similarity as a common basis for categorization and recognition memory. Journal of Neuroscience. 2014; 34(22):7472–7484. [PubMed: 24872552]

Dayan P, Abbott LF. Theoretical neuroscience: Computational and mathematical modeling of neural systems. Cambridge, MA: MIT Press; 2005.

Erlhagen W, Schöner G. Dynamic field theory of movement preparation. Psychological Review. 2002; 109(3):545–572. [PubMed: 12088245]

Enserink M, Kupferschmidt K. European neuroscientists revolt against the E.U.'s Human Brain Project. Science. 2014

Forstmann BU, Ratcliff R, Wagenmakers EJ. Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. Annual Review of Psychology. 2016; 67:641–666.

Forstmann BU, Wagenmakers EJ, Eichele T, Brown S, Serences JT. Reciprocal relations between cognitive neuroscience and formal cognitive models: Opposites attract? Trends in Cognitive Sciences. 2011; 15:272–279. [PubMed: 21612972]

Forstmann BU, Wagenmakers E-J, editorsAn introduction to model-based cognitive neuroscience. Springer; 2015.

Gluth S, Rieskamp J. Variability in behavior that cognitive models do not explain can be linked to neuroimaging data. Journal of Mathematical Psychology. 2017; 76:104–116.

Gold JI, Shadlen MN. The neural basis of decision making. Annual Review of Neuroscience. 2007; 30:535–574.

Grice GR. Stimulus intensity and response evocation. Psychological Review. 1968; 75(5):359–373. [PubMed: 4879423]

Gureckis TM, Love BC. Common mechanisms in infant and adult category learning. Infancy. 2004; 5:173–198.

Hanes DP, Schall JD. Neural control of voluntary movement initiation. Science. 1996; 274:427–430. [PubMed: 8832893]

Hawkins GE, Mittner M, Forstmann BU, Heathcote A. On the efficiency of neurally-informed cognitive models to identify latent cognitive states. Journal of Mathematical Psychology. 2017; 76:142–155.

Ide JS, Shenoy P, Yu AJ, Li CS. Bayesian prediction and evaluation in the anterior cingulate cortex. Journal of Neuroscience. 2013; 33(5):2039–2047. [PubMed: 23365241]

Khaligh-Razavi SM, Henriksson L, Kay K, Kriegeskorte N. Fixed versus mixed RSA: Explaining visual representations by fixed and mixed feature sets from shallow and deep computational models. Journal of Mathematical Psychology. 2017; 76:184–197. [PubMed: 28298702]

Khaligh-Razavi SM, Kriegeskorte N. Deep supervised, but not unsupervised, models may explain IT cortical representation. PLoS Computational Biology. 2014; 10:e1003915. [PubMed: 25375136]

Kragel JE, Morton NW, Polyn SM. Neural activity in the medial temporal lobe reveals the fidelity of mental time travel. Journal of Neuroscience. 2015; 35(7):2914–2926. [PubMed: 25698731]

Kriegeskorte N. Relating population-code representations between man, monkey, and computational models. Frontiers in Neuroscience. 2009; 3:363–373. [PubMed: 20198153]

Kriegeskorte N. Deep neural networks: A new framework for modeling biological vision and brain information processing. Annual Review of Vision Science. 2015; 1:417–446.

LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015; 521(7553):436–444. [PubMed: 26017442]

Logan GD, Yamaguchi M, Schall GD, Palmeri TJ. Inhibitory control in mind and brain 2.0: A blocked-input model of saccadic countermanding. Psychological Review. 2015; 122:115–147. [PubMed: 25706403]

Love BC. The algorithmic level is the bridge between computation and brain. Topics in Cognitive Science. 2015; 7(2):230–242. [PubMed: 25823496]

Love BC, Gureckis TM. Cognitive, Affective, & Behavioral Neuroscience. 2007. Models in search of a brain; 90–108.

Love BC, Medin DL, Gureckis TM. SUSTAIN: A network model of category learning. Psychological Review. 2004; 111(2):309–331. [PubMed: 15065912]

Mack ML, Preston AR, Love BC. Decoding the brain's algorithm for categorization from its neural implementation. Current Biology. 2013; 23:2023–2027. [PubMed: 24094852]

Mack ML, Preston AR, Love BC. Dynamic updating of hippocampal object representations reflects new conceptual knowledge. Proceedings of the National Academy of Sciences. 2016 in press.

Marr D. Vision: A computational investigation into the human representation and processing of visual information. New York: Freeman; 1982.

Mazurek ME, Roitman JD, Ditterich J, Shadlen MN. A role for neural integrators in perceptual decision making. Cerebral Cortex. 2003; 13:1257–1269. [PubMed: 14576217]

Mittner M, Boekel W, Tucker AM, Turner BM, Heathcote A, Forstmann BU. When the brain takes a break: A model-based analysis of mind wandering. Journal of Neuroscience. 2014; 34:16286–16295. [PubMed: 25471568]

Nosofsky RM, Palmeri TJ. An exemplar-based random walk model of speeded classification. Psychological Review. 1997; 104:266–299. [PubMed: 9127583]

Nunez MD, Vandekerckhove J, Srinivasan R. How attention influences perceptual decision making: Single-trial EEG correlates of drift-diffusion model parameters. Journal of Mathematical Psychology. 2017; 76:117–130. [PubMed: 28435173]

Oaksford M, Chater N. Bayesian rationality: The probabilistic approach to human reasoning. Oxford University Press; 2007.

Palmeri TJ. Exemplar similarity and the development of automaticity. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1997; 23:324–354.

Palmeri TJ. An exemplar of model-based cognitive neuroscience. Trends in Cognitive Science. 2014; 18(2):67–69.

Palmeri TJ, Schall JD, Logan GD. Neurocognitive modeling of perceptual decision making. In: Busemeyer JR, Townsend J, Wang ZJ, Eidels A, editorsOxford handbook of computational and mathematical psychology. Oxford University Press; 2015.

Pratte MS, Ling S, Swisher JD, Tong F. How attention extracts objects from noise. Journal of Neurophysiology. 2013; 100(6):1346–1356.

Pratte MS, Tong F. Integrating theoretical models with functional neuroimaging. Journal of Mathematical Psychology. 2017; 76:80–93. [PubMed: 28286346]

Purcell BA, Heitz RP, Cohen JY, Schall JD, Logan GD, Palmeri TJ. Neurally-constrained modeling of perceptual decision making. Psychological Review. 2010; 117:1113–1143. [PubMed: 20822291]

Purcell BA, Palmeri TJ. Relating accumulator model parameters and neural dynamics. Journal of Mathematical Psychology. 2017; 76:156–171. [PubMed: 28392584]

Purcell BA, Schall JD, Logan GD, Palmeri TJ. From salience to saccades: Multiple-alternative gated stochastic accumulator model of visual search. Journal of Neuroscience. 2012; 32(10):3433–3446. [PubMed: 22399766]

Ratcliff R, Cherian A, Segraves M. A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of simple two-choice decisions. Journal of Neurophysiology. 2003; 90:1392–1407. [PubMed: 12761282]

Ratcliff R, Rouder JN. Modeling response times for two-choice decisions. Psychological Science. 1998; 9:347–356.

Ratcliff R, Smith PL. A comparison of sequential sampling models for two-choice reaction time. Psychological Review. 2004; 111:333–367. [PubMed: 15065913]

Sakamoto Y, Love BC. Schematic influences on category learning and recognition memory. Journal of Experimental Psychology: General. 2004; 133:534–553. [PubMed: 15584805]

Schall JD. Neural basis of deciding, choosing and acting. Nature Reviews Neuroscience. 2001; 2:33–42. [PubMed: 11253357]

Schall JD. On building a bridge between brain and behavior. Annual Review of Psychology. 2004; 55:23–50.

Smith PL, Ratcliff R. Psychology and neurobiology of simple decisions. Trends in Neuroscience. 2004; 27:161–168.

Sun R, editorThe cambridge handbook of computational psychology. Cambridge University Press; 2008.

Teller DY. Linking propositions. Vision Research. 1984; 24:1233–1246. [PubMed: 6395480]

Tenenbaum JB, Kemp C, Griffiths TL, Goodman ND. How to grow a mind: Statistics, structure, and abstraction. Science. 2011; 331(6022):1279–1285. [PubMed: 21393536]

Theil S. Why the human brain project went wrong—and how to fix it. Scientific American. 2015; 313

Turner BM. Constraining cognitive abstractions through Bayesian modeling. In: Forstmann BU, Wagenmakers E-J, editorsAn introduction to model-based cognitive neuroscience. New York: Springer; 2015. 199–220.

Turner BM, Forstmann BU, Love B, Palmeri TJ, Van Maanen L. Approaches to analysis in model-based cognitive neuroscience. Journal of Mathematical Psychology. 2017; 76:65–79.

Turner BM, Forstmann BU, Wagenmakers EJ, Brown SD, Sederberg PB, Steyvers M. A Bayesian framework for simultaneously modeling neural and behavioral data. NeuroImage. 2013; 72:193–206. [PubMed: 23370060]

Turner BM, Van Maanen L, Forstmann BU. Combining cognitive abstractions with neurophysiology: The neural drift diffusion model. Psychological Review. 2015; 122:312–336. [PubMed: 25844875]

van Gerven MAJ. A primer on encoding models in sensory neuroscience. Journal of Mathematical Psychology. 2017; 76:172–183.

van Ravenzwaaij D, Provost A, Brown SD. A confirmatory approach for integrating neural and behavioral data into a single model. Journal of Mathematical Psychology. 2017; 76:131–141.

Wijeakumar S, Ambrose JP, Spencer JP, Curtu R. Model-based functional neuroimaging using dynamic neural fields: An integrative cognitive neuroscience approach. Journal of Mathematical Psychology. 2017; 76:212–235. [PubMed: 29118459]

Yamins DL, Hong H, Cadieu CF, Solomon EA, Seibert D, DiCarlo JJ. Performance-optimized hierarchical models predict neural responses in higher visual cortex. Proceedings of the National Academy of Sciences. 2014; 111(23):8619–8624.

Zandbelt BB, Purcell BA, Palmeri TJ, Logan GD, Schall JD. Response times from ensembles of accumulators. Proceedings of the National Academy of Sciences. 2014; 111(7):2848–2853.

HIGHLIGHTS

- Mathematical psychology and cognitive neuroscience come together in a powerful new approach called *model-based cognitive neuroscience*.
- This approach can both inform cognitive modeling and help to interpret neural measures.
- This article provides an introduction to the field of model-based cognitive neuroscience and to the articles contained within this special issue.