

Perspective

A Healthy Fear of the Unknown: Perspectives on the Interpretation of Parameter Fits from Computational Models in Neuroscience

Matthew R. Nassar, Joshua I. Gold*

Department of Neuroscience, University of Pennsylvania, Philadelphia, Pennsylvania, United States of America

Abstract: Fitting models to behavior is commonly used to infer the latent computational factors responsible for generating behavior. However, the complexity of many behaviors can handicap the interpretation of such models. Here we provide perspectives on problems that can arise when interpreting parameter fits from models that provide incomplete descriptions of behavior. We illustrate these problems by fitting commonly used and neurophysiologically motivated reinforcement-learning models to simulated behavioral data sets from learning tasks. These model fits can pass a host of standard goodness-of-fit tests and other model-selection diagnostics even when the models do not provide a complete description of the behavioral data. We show that such incomplete models can be misleading by yielding biased estimates of the parameters explicitly included in the models. This problem is particularly pernicious when the neglected factors are unknown and therefore not easily identified by model comparisons and similar methods. An obvious conclusion is that a parsimonious description of behavioral data does not necessarily imply an accurate description of the underlying computations. Moreover, general goodness-of-fit measures are not a strong basis to support claims that a particular model can provide a generalized understanding of the computations that govern behavior. To help overcome these challenges, we advocate the design of tasks that provide direct reports of the computational variables of interest. Such direct reports complement model-fitting approaches by providing a more complete, albeit possibly more task-specific, representation of the factors that drive behavior. Computational models then provide a means to connect such task-specific results to a more general algorithmic understanding of the brain.

The use of models to infer the neural computations that underlie behavior is becoming increasingly common in neuroscience research, especially for cognitive and perceptual tasks involving decision making and learning. As their sophistication and usefulness expand, these models become increasingly central to the design, analysis, and interpretation of experiments. We consider this development to be generally positive but provide here some perspectives on the challenges inherent to this approach, particularly when behavior might be driven by unexpected factors that can complicate the interpretation of model fits. Our goal is to raise awareness of these issues and present complementary approaches that can help ensure that our understanding of the brain does not become overly conditioned to the quality of existing models fit to particular data sets.

We illustrate these challenges using a set of models that describe the ongoing process of learning values to guide actions and that are used extensively in the field of cognitive neuroscience [1–13]. These models adjust expectations about future outcomes according to the difference between actual and predicted outcomes, known as the prediction error. Originally developed in parallel in both animal- and machine-learning fields [14–16], this relatively simple form of reinforcement-learning algorithm (often referred to as a “delta rule” because the prediction error is typically represented by the Greek symbol delta (δ) in the equations) has: 1) provided

efficient solutions to a broad array of biologically relevant problems [15]; 2) accounted for many, but not all, learning phenomena exhibited by both human and nonhuman subjects [17,18]; 3) provided a generative architecture that has been used to predict behavior across tasks, compare brain activity to learning variables within a single task, and explore the range of possible behaviors that one might expect to find in a variable population [19,20]; and 4) guided an understanding of the neural computations expressed by the brainstem dopaminergic system [21]. These successes have led to the proposal that the interpretation of delta-rule model parameters fit to behavioral data from human subjects performing simple learning tasks might serve as a more precise diagnostic tool for certain mental disorders than existing methods [22–24]. Thus reinforcement-learning models are becoming highly influential in guiding and filtering our understanding of normal and pathological brain function.

Here we focus on the interpretation of a term in most delta-rule models called the learning rate. The learning rate, α , determines the amount of influence that the prediction error, δ , associated with a given outcome has on the new expectation of future outcomes, E :

$$E_{t+1} = E_t + \alpha \times \delta_{t+1} \quad (\text{EQ1})$$

As its name implies, the learning rate determines how quickly the model adapts

Citation: Nassar MR, Gold JI (2013) A Healthy Fear of the Unknown: Perspectives on the Interpretation of Parameter Fits from Computational Models in Neuroscience. *PLoS Comput Biol* 9(4): e1003015. doi:10.1371/journal.pcbi.1003015

Editor: Olaf Sporns, Indiana University, United States of America

Published: April 4, 2013

Copyright: © 2013 Nassar, Gold. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by NIH EY015260 and MH098899, the McKnight Endowment Fund for Neuroscience, the Burroughs-Wellcome Fund, the Sloan Foundation, and a Ruth L. Kirschstein National Research Service Award MH093099 (MN). The funders had no role in the preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: jigold@mail.med.upenn.edu

to errors. A fixed value near zero implies that expectations are updated slowly, essentially averaging over a long history of past outcomes. In contrast, a fixed value near one implies that expectations are updated quickly to match the most recent outcomes. Thus, the learning rate can be interpreted as the amount of influence each unpredicted outcome exerts on the subsequent expectation. These updated expectations can, in turn, be used to select actions, often through a soft-max function with an inverse-temperature parameter. This parameter can be adjusted to optimize the trade-off between exploiting actions known to be valuable in the present (emphasized at higher inverse temperatures) and exploring actions that might be valuable in the future (emphasized at lower inverse temperatures) [12,13,15,25].

Recent work has highlighted the advantages of using learning rates that, instead of remaining fixed, are adjusted adaptively according to environmental dynamics [9–11,26–28]. For example, adaptive learning rates can help ensure that expectations remain relatively stable in stationary environments but change rapidly in response to abrupt environmental changes. Consistent with this idea, human behavior in tasks containing abrupt changes conforms to models in which the influence of each outcome depends on the statistics of other recent outcomes. Such rational adjustments of learning rate are most prominent after changes in action-outcome contingencies that lead to surprisingly large prediction errors [9–11].

Here we consider in detail two of these change-point tasks. The first, an estimation task, requires subjects to predict the next in a series of outcomes (randomly generated numbers) [9]. Each outcome is drawn from a normal distribution with a fixed mean and variance. However, the mean of this distribution is occasionally reset at random times, producing abrupt change-points in the series of outcomes. Learning rates can be measured directly on a trial-by-trial basis, using predictions and outcomes plugged into Eq. 1. Previous work showed that subjects performing this task tended to use learning rates that were consistent with predictions from a reduced form of a Bayesian ideal-observer algorithm, including a positive relationship between error magnitude and learning rate. However, the details of this relationship varied considerably across individual subjects. Some subjects tended to use highly adaptive learning rates, including values near zero following small errors and values near one following

surprisingly large prediction errors. In contrast, other subjects used a much narrower range of learning rates, choosing similar values over most conditions. This across-subject variability was described by a flexible model that could generate behaviors ranging from that of a fixed learning-rate delta rule to that of the reduced Bayesian algorithm, depending on the value of a learning rate “adaptive-ness” parameter.

The second task is a four-alternative forced-choice task that includes occasional, un signaled change-points in the probabilistic associations of monetary rewards for each choice target [11]. Learning rates are not measured directly, as in the estimation task, but rather inferred from model fits. The best-fitting models incorporate learning rates that increase transiently after unexpectedly large errors, although the magnitude of this increase differs across subjects. The existence of this kind of across-subject variability can have dramatic effects on the interpretation of best-fitting parameters from models that do not account for this variability explicitly. Here we illustrate this problem by fitting behavioral data corresponding to different forms of adaptive learning with delta-rule models that neglect adaptive learning entirely. However, we emphasize that this problem is not limited to adaptive learning but can also arise when neglecting other factors that can influence performance on learning tasks, such as a tendency to repeat choices [29,30], and more generally whenever oversimplified models are fit to complex behavioral data.

We used simulations of the two tasks to illustrate how fitting models with fixed learning rates to behavior that is based on adaptive learning rates can lead to misleading conclusions. For each task, behavioral data were simulated using a delta-rule inference algorithm with different levels of learning-rate adaptiveness coupled with a soft-max function for action selection. These simulated data were then fit, using maximum-likelihood methods, with a simpler model that included two free parameters: a fixed learning rate and the inverse temperature of a soft-max action-selection process (see Text S1). In all cases, the simpler, fixed learning-rate model was preferred over a null model constituting random choice behavior, even after penalizing for additional complexity (e.g., using BIC or AIC; see Text S1). Despite passing these model-selection criteria, we highlight two misleading conclusions that might be drawn from these fits: biased estimates of adaptive learning and of exploratory behavior.

The problem of misestimating adaptive learning is depicted in Figure 1A & B. Panel A shows simulations based on the estimation task. For this task, learning rate is measured directly as the proportion of the current prediction error used to update from the current prediction to the next prediction [9]. As expected, variability in measured learning rates tended to increase with learning-rate adaptiveness. The average value of measured learning rates also tended to decrease with learning-rate adaptiveness, because change-points that dictate high values of adaptive learning rates were relatively rare in our simulated tasks (black circles and error bars reflect median and interquartile range, respectively, across 800 simulated trials).

The model fits, however, tell a different story. When behavior was simulated using a fixed learning rate (learning-rate adaptiveness = 0), the best-fitting models naturally captured the appropriate value. However, when behavior was simulated using increasingly adaptive learning rates, the fixed learning-rate models returned systematically larger estimates of learning rate than were actually used by the simulated subjects (Figure 1A, gray points).

Panel B shows simulations based on the four-choice task, for which we determined the learning rate on each trial based on its value in the internal, generative process used in the simulations. Data from this task tell a similar story. Simulated learning rates were lower but more variable for more adaptive models (black circles and error bars reflect median and interquartile range), yet fit learning rates were higher for these same models (Figure 1B, gray points). These data suggest that periods of rapid learning (i.e., following change-points) are more influential than periods of slow learning on maximum-likelihood fits of the fixed learning-rate parameter, which thus becomes biased upwards when the underlying learning rate is adaptive.

The problem of misestimating exploratory behavior is depicted in Figure 1C & D. We first simulated behavior on both the estimation task and the four-choice task using a fixed learning rate and an action-selection process governed by an inverse-temperature parameter. In each case, fits from a model with a fixed learning rate and an inverse-temperature process returned appropriate estimates of the inverse temperature used in the generative process (left-most circles in Figure 1C & D, corresponding to learning-rate adaptiveness = 0).

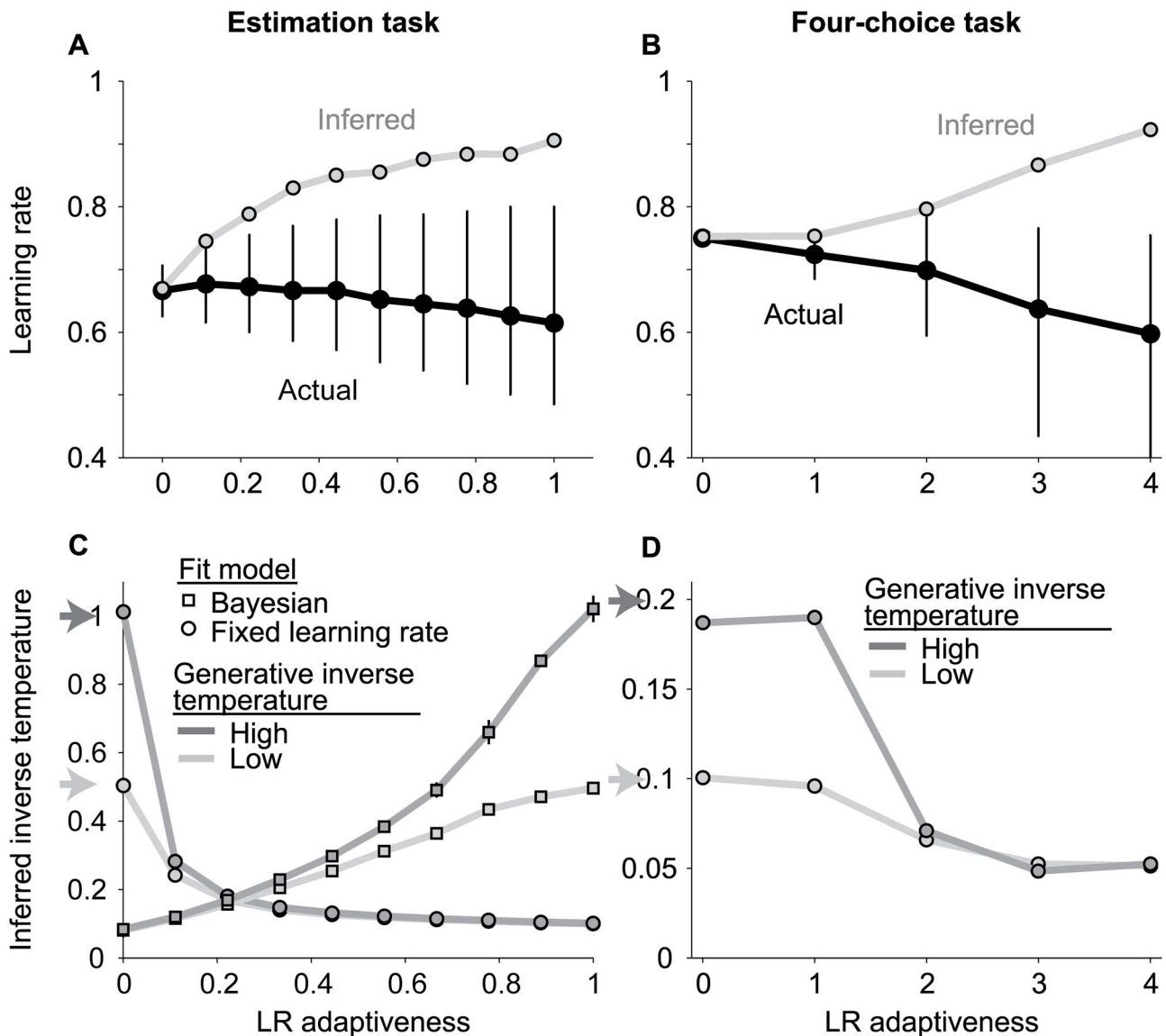


Figure 1. Learning-rate adaptiveness can be misinterpreted as elevated fixed learning rates and decreased inverse temperatures for the estimation (A,C) or four-alternative (B,D) tasks (see text). In all panels, the abscissa represents learning-rate adaptiveness (0 is equivalent to using a fixed learning rate; higher numbers indicate higher adaptiveness to unexpected errors). A & B. Actual (black) and model-inferred (gray) learning rates used by agents with different levels of learning-rate adaptiveness. Points and error bars represent the median and interquartile range, respectively, of data from six simulated sessions. C & D. Best-fitting values of the inverse-temperature parameter, intended to describe exploratory behavior, inferred using a fixed delta-rule (circles) or approximately Bayesian (squares) model. Shades of gray indicate the level of exploratory behavior of the simulated agent, as indicated. Arrows indicate the actual value of the inverse-temperature parameter used in the generative process. Points and error bars (obscured) represent the mean and standard error of the mean, respectively, of data from six simulated sessions.
doi:10.1371/journal.pcbi.1003015.g001

However, when the simulated subjects used increasingly adaptive learning rates, inverse-temperature fits from a fixed learning-rate model substantially overestimated the true variability in action selection (circles in Figure 1C & D: inferred inverse temperature decreases as learning-rate adaptiveness increases). Such biased parameter estimates were not simply a problem with the fixed learning-rate model. Fitting an alternative model

that used optimal (maximally adaptive) learning rates [9,31] to the behavior of the same simulated subjects yielded a complementary pattern of biases: the model accurately inferred the level of exploratory action selection for simulated subjects that choose learning rates adaptively but overestimated this quantity for subjects that used simpler strategies of less-adaptive, or even fixed, learning rates (squares in Figure 1C: inferred inverse temperature

decreases as learning-rate adaptiveness decreases). For both models, these problems were not apparent from standard analyses of best-fitting parameter values, which had similar confidence intervals and covariance estimates for biased and unbiased fit conditions (see Text S1). These problems also did not simply reflect difficulties in estimating model parameters when the inverse temperature was low and behavior was more random, because the

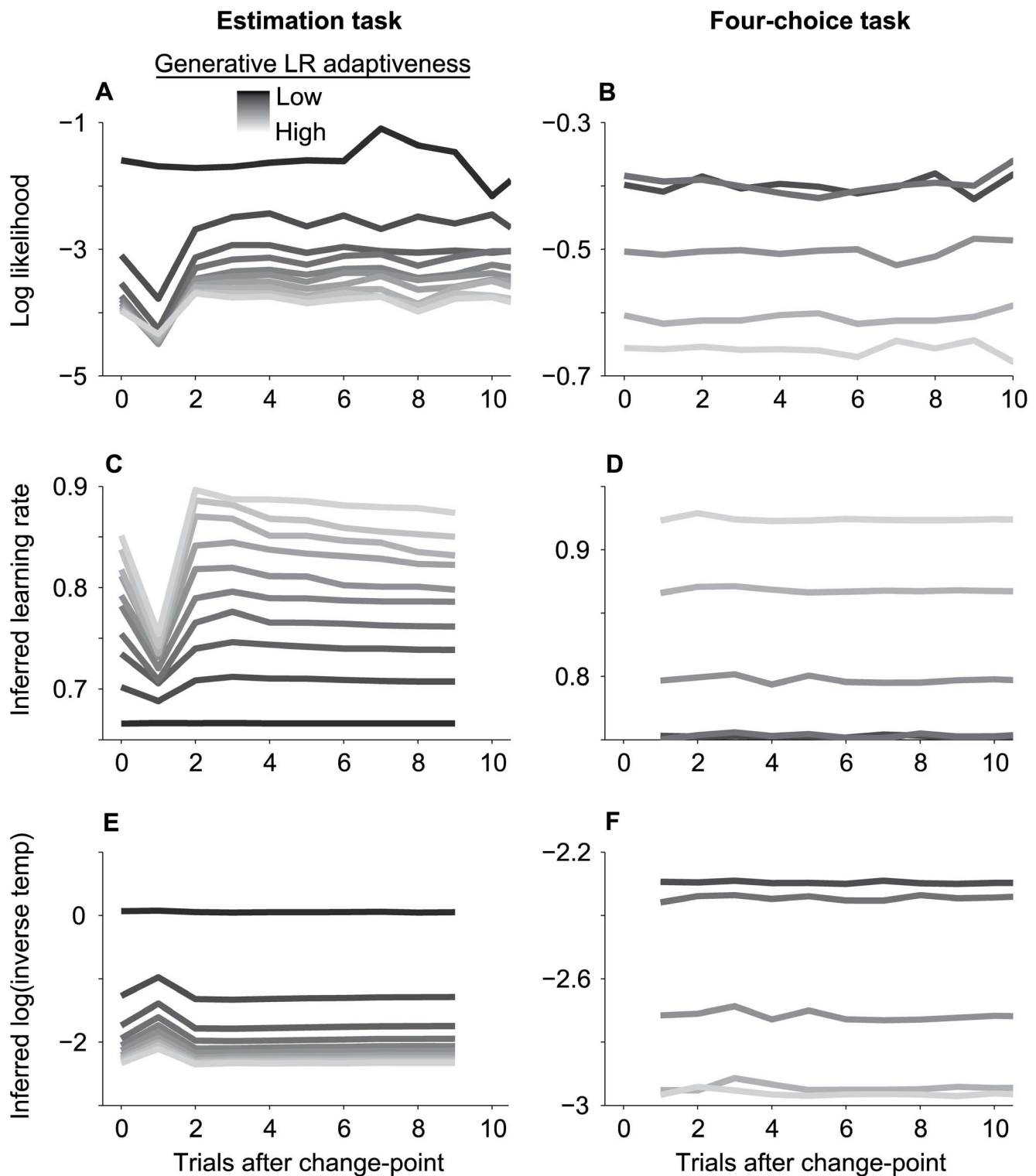


Figure 2. Poor fits from models that ignore learning-rate adaptiveness are easily identified in the estimation, but not the four-choice, task. A & B. Mean log-likelihood associated with a fixed learning-rate model, per simulated trial from the estimation (A) or four-choice (B) task, aligned to change-points in the generative process. Lighter shades of gray represent data from simulated agents with higher levels of learning-rate adaptiveness. C–F. Learning rates (C & D) or inverse temperatures (E & F) inferred from model fits that exclude log-likelihood information from trials occurring 0–10 trials after change-points (abscissa) for estimation (C & E) and four-choice (D & F) tasks. The transient changes in A, C, and E evident for all but the least adaptive simulated agents reflect the fixed learning-rate model’s inability to account for behavior just following change-points on the estimation task; no comparable effects are evident for the four-choice task.
doi:10.1371/journal.pcbi.1003015.g002

problem was also apparent when the inverse temperature was high. Thus, subtle differences in learning that were not accounted for by the inference model caused underestimation of the inverse-temperature parameter, which might be misinterpreted as increases in exploratory action selection.

Diagnosing these kinds of problems is difficult, especially when the subtle aspect of behavior that is missing from the model is unknown. Model-selection practices that compare likelihoods of various models (after either cross validation or penalization of parameter numbers) are useful for identifying the better of two or more models with respect to particular data sets. However, these practices require a priori knowledge of the models to be tested and cannot, by themselves, indicate what might be missing from the tested models. One might be tempted to interpret likelihoods directly and set a criterion for what might be considered a “good” model. However, these metrics cannot say whether or not a model is correct (or even sufficiently good, given that no fit model is truly correct). For example, consider a test of the suitability of a fixed learning-rate model for simulated subjects that can vary in terms of learning-rate adaptiveness and exploratory behavior. Similar values of AIC, BIC, and other likelihood-based quantities are obtained for fixed delta-rule models fit to two very different subjects: one who uses a fixed learning rate, which is consistent with the model, and relatively high exploration; and another who uses a highly adaptive learning rate, which is inconsistent with the model, and relatively low exploration. Interpretation of parameter fits from the latter case would be misleading, whereas parameter fits from the former would be asymptotically unbiased and thus more informative.

To overcome these limitations, it is sometimes effective to look for indications that a model is failing under specific sets of conditions for which behavior is heavily influenced by the assumptions of the model. For the case of adaptive learning, fixed learning-rate models fail to address adaptive responses to inferred change-points in the action-outcome contingency. Thus, it can be instructive to examine the likelihoods of these models computed for choice data collected shortly after change-points. For the case of the estimation task, a fixed learning-rate model shows an obvious inability to account for data from trials just after a change-point for all but the least adaptive simulated subjects (Figure 2A; dip in log-likelihood at trial

1). However, this approach is not effective for the four-choice task (Figure 2B).

Another potentially useful approach for diagnosing misleading parameter fits is to compute these fits using subsets of data that might correspond to different best-fitting values of certain parameters. For the estimation task, eliminating data from trials immediately following change-points has dramatic effects on fits for both learning rate (Figure 2C) and inverse temperature (Figure 2E). However, this diagnostic approach is far less effective for the four-choice task, for which adjustments in learning rate occur with a longer and more variable time course following change-points (Figure 2D & F). Thus, for tasks like the estimation task that provide explicit information about the subject’s underlying expectations, the insufficiency of the fixed learning-rate model can be fairly simple to diagnose. However, for tasks like the four-choice task in which information about the subject’s expectations is limited to inferences based on less-informative choice behavior, parameter biases are still large (Figure 1B & D), but model insufficiency is far less apparent.

A sobering conclusion that can be drawn from these examples is that even when the parameter fits from a computational model are reasonably likely to produce a data set, and even when this likelihood is robust to perturbations in the specific trials that are fit or the settings of other parameters in the model, the model might still be missing specific features of the data. Missing even a fairly nuanced feature of the data (such as adaptive learning) can lead the parameters in the model to account for the feature in surprising ways. These unexpected influences can lead to parameter fits that, if interpreted naïvely, might suggest computational relationships that are unrelated to, or even opposite to, the true underlying relationships. Here we use an example from reinforcement learning, but the lessons apply to any model-fitting procedure that requires the interpretation of best-fitting parameter values. Certain parameters, like the inverse-temperature parameter in reinforcement-learning models, are particularly susceptible to this problem, because they are always sensitive to other sources of behavioral variability that are incompletely described by the rest of the model.

These challenges highlight the narrow wire on which the computational neuroscientist walks. On one hand, we seek to generalize a wide array of physiological and behavioral data from different tasks onto a tractable set

of computational principles. On the other hand, the results that we obtain from each experiment are conditioned on assumptions from the particular model through which they are obtained. We believe that the goals of computational neuroscience are possible even in the face of this contradiction. Obtaining generalizable results depends on not only good modeling practices [32] but also the extensive use of descriptive statistics to dissect and interpret data from both experiments and simulated model data. For example, the estimation task described above was designed to allow learning rates from individual trials to be computed directly and not inferred via model fits to resulting choice behaviors. This approach revealed clear task-dependent effects on adaptive learning [9]. In principle, congruence between these kinds of direct analyses of behavioral data and fit model parameters can help support interpretations of those parameters and has the advantage of testing modeling assumptions and predictions explicitly rather than via comparisons of different model sets [8,33,34]. In contrast, inconsistencies between direct analyses and fit model parameters can help guide how the model can be modified or expanded—keeping in mind, of course, that adding to a model’s complexity can improve its overall fit to the data but often by overfitting to spurious features of the data and making it more difficult to interpret the contributions of individual parameters [35].

In summary, model fits to behavioral data can provide useful and important insights into the neurocomputational principles underlying such behavior but should not replace good experimental designs that explicitly isolate, manipulate, and/or measure the behavioral processes of interest. Combining such designs with both model fitting and other kinds of analyses can support steady progress in attaining a more general understanding of the neural basis for complex behaviors that is not overly tied to a particular model or behavioral test.

Supporting Information

Text S1 Provides methods for simulations and model fitting as well as Bayesian information criterion values for each set of models. (DOCX)

Acknowledgments

We thank David Brainard, Peter Dayan, Long Ding, Yin Li, Mike Shadlen, Michael Todd, Takahiro Doi, and Robert Wilson for helpful comments.

References

1. Beeler JA, Daw N, Frazier CRM, Zhuang X (2010) Tonic dopamine modulates exploitation of reward learning. *Front Behav Neurosci* 4: 170.
2. Doll BB, Hutchison KE, Frank MJ (2011) Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *J Neurosci* 31: 6188–6198.
3. Frank MJ, Doll BB, Oas-Terpstra J, Moreno F (2009) Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nat Neurosci* 12: 1062–1068.
4. Jepma M, Nieuwenhuis S (2011) Pupil diameter predicts changes in the exploration-exploitation trade-off: evidence for the adaptive gain theory. *J Cogn Neurosci* 23: 1587–1596.
5. Seo H, Lee D (2008) Cortical mechanisms for reinforcement learning in competitive games. *Philos Trans R Soc Lond B Biol Sci* 363: 3845–3857.
6. Strauss GP, Frank MJ, Waltz JA, Kasanova Z, Herbener ES, et al. (2011) Deficits in positive reinforcement learning and uncertainty-driven exploration are associated with distinct aspects of negative symptoms in schizophrenia. *Biol Psychiatry* 69: 424–431.
7. Sul JH, Jo S, Lee D, Jung MW (2011) Role of rodent secondary motor cortex in value-based action selection. *Nat Neurosci* 14: 1202–1208.
8. Walton ME, Behrens TEJ, Buckley MJ, Rushworth MF (2010) Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron* 65: 927–939.
9. Nassar MR, Wilson RC, Heasly B, Gold JI (2010) An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J Neurosci* 30: 12366–12378.
10. Behrens TE, Woolrich MW, Walton ME, Rushworth MF (2007) Learning the value of information in an uncertain world. *Nat Neurosci* 10: 1214–1221.
11. Krugel LK, Biele G, Mohr PNC, Li S-C, Heekeren HR (2009) Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proc Natl Acad Sci U S A* 106: 17951–17956.
12. Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. *Nature* 441: 876–879.
13. Luksys G, Gerstner W, Sandi C (2009) Stress, genotype and norepinephrine in the prediction of mouse behavior using reinforcement learning. *Nat Neurosci* 12: 1180–1186.
14. Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black AH, Prokasy WF, editors. *Classical conditioning II: current research and theory*. New York: Appleton-Century-Crofts. pp. 64–99.
15. Sutton RS, Barto AG (1998) *Reinforcement learning: an introduction*. Cambridge (Massachusetts): MIT Press. 342 p.
16. Bertsekas DP, Tsitsiklis JN (1996) *Neuro-dynamic programming*. Belmont (Massachusetts): Athena Scientific. 512 p.
17. Dayan P, Abbott LF (2001) *Theoretical neuroscience*. Cambridge (Massachusetts): MIT Press.
18. Shanks DR (1995) *The psychology of associative learning (problems in the behavioral sciences)*. Gray J, editor. Cambridge: Cambridge University Press. 194 p.
19. Miller RR, Barnet RC, Grahame NJ (1995) Assessment of the Rescorla-Wagner model. *Psychol Bull* 117: 363.
20. Dayan P, Daw N (2008) Decision theory, reinforcement learning, and the brain. *Cogn Affect Behav Neurosci* 8: 429–453.
21. Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275: 1593–1599.
22. Huys QJM, Moutoussis M, Williams J (2011) Are computational models of any use to psychiatry? *Neural Netw* 24: 544–551.
23. Huys Q, Vogelstein J, Dayan P (2009) Psychiatry: insights into depression through normative decision-making models. *Adv Neural Inf Process Syst* 21: 729–736.
24. Maia TV, Frank MJ (2011) From reinforcement learning models to psychiatric and neurological disorders. *Nat Neurosci* 14: 154–162.
25. Ishii S, Yoshida W, Yoshimoto J (2002) Control of exploitation-exploration meta-parameter in reinforcement learning. *Neural Netw* 15: 665–687.
26. Yu AJ, Dayan P (2005) Uncertainty, neuromodulation, and attention. *Neuron* 46: 681–692.
27. Mathys C, Daunizeau J, Friston KJ, Stephan KE (2011) A bayesian foundation for individual learning under uncertainty. *Front Hum Neurosci* 5: 39.
28. Preusschoff K, Bossaerts P (2007) Adding prediction risk to the theory of reward learning. *Ann N Y Acad Sci* 1104: 135–146.
29. Lau B, Glimcher PW (2005) Dynamic response-by-response models of matching behavior in rhesus monkeys. *J Exp Anal Behav* 84: 555–579.
30. Schönberg T, Daw ND, Joel D, O'Doherty JP (2007) Reinforcement learning signals in the human striatum distinguish learners from non-learners during reward-based decision making. *J Neurosci* 27: 12860–12867.
31. Wilson RC, Nassar MR, Gold JI (2010) Bayesian online learning of the hazard rate in change-point problems. *Neural Comput* 22: 2452–2476.
32. Daw ND (2009) Trial-by-trial data analysis using computational models. In: Phelps E, Robbins T, Delgado M, editors. *Decision making, affect, and learning: attention and performance XXIII*. Oxford: Oxford University Press. pp. 3–38.
33. Ding L, Gold JI (2012) Separate, causal roles of the caudate in saccadic choice and execution in a perceptual decision task. *Neuron* 75: 865–874.
34. Frank MJ, Seeberger LC, O'Reilly RC (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306: 1940–1943.
35. Pitt MA, Myung IJ (2002) When a good fit can be bad. *Trends Cogn Sci* 6: 421–425.