Journal of Motor Behavior, Vol. 42, No. 6, 2010

RESEARCH ARTICLE

In Praise of "False" Models and Rich Data

Hugo L. Fernandes[1,2], Konrad P. Kording[1,2]

[1]Department of Physiology, Physical Medicine and Rehabilitation, Northwestern University and Rehabilitation Institute of Chicago, Illinois. [2]Department of Applied Mathematics, Northwestern University and Rehabilitation Institute of Chicago, Illinois.

ABSTRACT. The authors argue that "true" models that aim at faithfully mimicking or reproducing every property of the sensorimotor system cannot be compact as they need many free parameters. Consequently, most scientists in motor control use what are called "false" models—models that derive from well-defined approximations. The authors conceptualize these models as a priori limited in scope and approximate. As such, they argue that a quantitative characterization of the deviations between the system and the model, more than the mere act of falsifying, allows scientists to make progress in understanding the sensorimotor system. Ultimately, this process should result in models that explain as much data variance as possible. The authors conclude by arguing that progress in that direction could strongly benefit from databases of experimental results and collections of models.

Keywords: Bayesianism, falsifiability, motor control

Among other books, we own typical textbooks introducing the principles in physics (Feynman, Leighton, & Sands, 1963) and neuroscience (Kandel, Schwartz, & Jessell, 2000). We start our reflections on falsification comparing the models of motor control suggested in the neuroscience textbook with the models of mechanics suggested in the physics textbook.[1] The models[2] in the book of physics are very compact[3]—and we find everything we may need to simulate countless real-life mechanical situations in the introductory book. For example, calculating the apple's fall from the tree can be described to near arbitrary precision by a simple differential equation with one free parameter that can be easily solved. The chapter on motor control, on the other hand, describes a large number of known neural phenomena. However, it does not provide nearly enough information to simulate the movement of a hand as it captures the apple. Although the chapter exposes quite a bit of the complexity of how the nervous system controls movement, it does not provide the necessary parameters to actually simulate the system. Many phenomena of real-world importance in physics can be modeled precisely by a compact model that can be explained in the space afforded by a textbook. This cannot necessarily be said about models of motor control.

This difference in compactness between models may have different reasons. It may be that models of motor control are not yet compact because the field is relatively new (Kuhn, 1970). However, it may also be that there is a fundamental difference between the compactness of the kinds of models researchers need to describe effects of real-world importance. In models of motor control many parameters may be necessary to capture the already countless demonstrated influences of the environment on humans' behavior. Many additional parameters may be necessary to capture the influence of the genome. The working hypothesis of this paper is that models of motor control can only either precisely explain the bulk of the data or be compact, but not both at the same time.[4]

Compactness is useful in many ways: Scientists can only think about, communicate, efficiently simulate, and constrain the parameters of compact models. Therefore, compact models have traditionally had exceptional influence on the design of new experiments. Scientists may ask which kinds of experiments and which kind of model development is most productive in a domain in which precise models are not compact.

In this article, we first discuss the kinds of models in motor control and their role as tools for research, focusing on the difference between models that aim at mimicking the nervous system in a detailed way (which we call "true" models) and other models that do not (which we call "false" models). We also distinguish between models that focus on implementations, algorithms, and computational objectives. We discuss falsification in the context of each of these kinds of models, and doing so, we argue that it is not falsification per se that is interesting but that the structure of differences between model predictions and experimental data should be scientists' main concern. Last, we discuss which implications our considerations have for improvement in the construction of models and experiments and argue that data and model sharing promises to accelerate progress in motor control.

Models

Intuitive and Mathematical Models

Before we discuss the implication of the difference between typical models in physics and motor control, we want to discuss the nature of models. Knowledge is only really useful when it generalizes, and a good experiment typically uncovers effects that do not just apply to the narrow range of situations addressed in the experiment. Models are devices that allow scientists to formalize knowledge and they aim at explaining data. All other things being equal, the best model is the one that explains more data variance[5] because it is thus a better approximation to reality. Scientists may distinguish between intuitive and mathematical models.

Correspondence address: Hugo L. Fernandes, Rehabilitation Institute of Chicago, Northwestern University, 345 E. Superior St., Attn: Onterie Rm 931, Chicago, IL 60611, USA. e-mail: hugoguh@gmail.com

In both model classes useful theories have been formulated. For example, Santiago Ramón y Cajal's neuron doctrine states that individual neurons give rise to the function of the nervous system. This intuitive model explains a lot of data about the nervous system. Hodgkin and Huxley's model for spike generation on the other hand is mathematically fully spelled out as a set of differential equations. It also explains a lot of data. Although the use of mathematical language allows a particularly precise definition of a model, most useful models in neuroscience derive from an intuitive description of data. The discussion in this paper is meant to apply equally to both kinds of models. We assume that much of the value of a model comes from the amount of data that it explains[6] and from its compactness.

"True" Bottom-Up Models

> In that Empire, the art of cartography attained such perfection that the map of a single province occupied the entirety of a city, and the map of the empire, the entirety of a province. In time, those unconscionable maps no longer satisfied, and the cartographers guilds struck a map of the empire whose size was that of the empire, and which coincided point for point with it. The following generations, who were not so fond of the study of cartography as their forebears had been, saw that that vast map was useless, and not without some pitilessness was it, that they delivered it up to the inclemencies of sun and winters. (Borges, 1999, p. 325)

Before discussing the role of models in motor control in more detail, we want to discuss one class of popular models: the kind of model that strives to mimic the nervous system as close as possible.[7] According to this philosophy, a good model should describe the brain in terms of the same units as they exist in the brain: synapses, neurons, connections, and brain areas. These models are bottom-up in the sense that behavior of neurons is to be explained in terms of synapses and behavior of brain areas is to be explained in terms of neurons.[8] The advantage of having such a model for medicine is obvious: scientists could test medical interventions. For example, if scientists had such a model, they could test which patterns of deep brain stimulation would optimally treat Parkinson's disease. In fact, it seems quite appropriate to call this a "true" model because it could in principle capture all aspects of neural processing and it may be argued that only this kind of a model could withstand a wide range of attempts at falsification.

Although this kind of model is clearly desirable to ask certain questions, there are various issues that limit its potential usefulness. We subsequently discuss these three issues with "true" models: (a) It seems to us that they can not be compact. (b) It appears that it is currently impossible to measure the necessary parameters to a level where complex phenomena such as behavior can be predicted. (c) Even if we had a fully working model of that kind (e.g., a brain in the computer) it could be argued that we do not really understand the underlying process.

First, it seems that a full specification of the nervous system must have many free parameters.[9] On top of the parameters that describe behavior, other parameters would be necessary to describe the dynamics of each neuron (e.g., channel densities, synaptic weights, dendritic arbors). If all of these parameters are to be specified then describing any significant part of the nervous system requires more information than could be written into a book.

There might be ways of simplifying the description (principles). For example, if neurons are actually connected randomly (Maass, Natschläger, & Markram, 2002) or alternatively if they all share the same properties it is easier to describe the ensemble.[10] If every neuron in a motor cortex would have cosine tuning properties (Georgopoulos, Schwartz, & Kettner, 1986), then scientists could describe them more compactly. However, in brain areas in which this same coding scheme model has been tested the near universal answer was that actually neurons are not that similar to one another. In general it looks like different neurons in the same brain area may have a wide spectrum of different properties (e.g., Morrow, Jordan, & Miller, 2007). Even in the primary visual cortex, probably the best studied brain area at present, a good part of the ongoing variance can not be explained by current models (Carandini et al., 2005) and it is possible to get the impression that this is a fundamental problem. Indeed, if the properties of the nervous system are obtained by learning mechanisms and learning is competitive (as in most present models), then the environment may actually drive neurons to be different from one another.[11]

Second, many experiments characterize properties of the nervous system, but it seems that taken together all these experiments cannot sufficiently constrain bottom-up models of the nervous system in all but the simplest cases. It is known in learning theory that virtually any nonlinear learning system has the universal approximation property (Kurt, Maxwell, & Halbert, 1990). If scientists have a system that is nonlinear and can learn, for example a neural network, then the system can approximate any input–output function to arbitrary precision as the number of degrees of freedom (e.g., neurons) goes to infinity. As such, the only way of properly constraining bottom-up models is the full characterization of essentially all their properties.

Third, let us imagine success in this way of modeling (Markram, 2006). Scientists have a completely faithful representation of the nervous system in a computer system and can do any simulated intervention on it. In this case, the program in the computer would be just as hard to understand as the brain. Obviously it would be possible to record from all neurons at the same time but even then, it is very difficult to understand a many-dimensional, highly nonlinear and complex system. Ultimately, a successful description of a system should explain the problem it solves, the algorithms it uses to solve these problems, and the implementation of this algorithm.

A Taxonomy of "False" Models

Because of the problems with "true" models previously discussed, it may be argued that a wide range of models is needed to lead to a convincing understanding of the nervous system. Marr (1982) introduced a taxonomy of three different kinds of models (see also Dayan & Abbott, 2001). According to Marr, it is possible to divide models into those that deal with the implementation of computation (Level 3), the algorithm used by the nervous system (Level 2), and the objective of computation (Level 1).[12] We subsequently argue that models at different levels demand fundamentally different experiments to put them to a test.

Scientists believe that there are advantages to the division into levels. Statements about an algorithm may be right even if the objective is wrong. Statements about the objective may be right if the algorithm is wrong. Obviously, for an algorithm to be believable there must be a possible implementation. However, for the same algorithm there may be countless possible implementations and vice-versa. The same holds true for the relationship between objectives and algorithms. This division thus may facilitate discussion about falsification.[13]

Marr Level 3 Models: Implementation

A range of models in motor control strives to understand the details of the implementation. Implementation models typically ask how aspects of the nervous system such as synapses, membrane potentials, and spikes work. For example, some models analyze how properties of motor neurons affect fluctuations of force output in muscles. Similarly, some models deal with the effect of muscle spindle feedback (Gielen & Houk, 1987) and the firing properties of spindle receptors (Prochazka & Gorassini, 1998). Some other models describe how single synapses may have more than one associated timescale (Fusi, Drew, & Abbott, 2005).

Marr Level 2 Models: Algorithm

Another range of models is concerned with algorithms. Algorithm in this context is typically related to the question of which computations are used by the nervous system. For example, depth may be estimated using the local crosscorrelation function across eyes. Other examples are the two timescale models of Smith, Ghazizadeh, and Shadmehr (2006) in which an algorithm of how the nervous system may learn is assumed: two learners systems adapt in parallel and the nervous system uses the sum of both learners for behavior. Yet another example is the algorithms described in the context of dynamical systems that explain many aspects of the dynamics of interlimb coordination (Haken, Kelso, & Bunz, 1985).

Marr Level 1 Models: Computational Objective

A third range of models asks about the objective of a computation. Computational objective in this context is typically related to the problem solved by the nervous system. A typical example is the question of if the nervous system combines cues from different modalities taking into account their uncertainty to obtain a minimum variance estimate (Kording, 2007). Other examples are wiring length models that ask how the nervous system may optimally wire itself (Chen, Hall, & Chklovskii, 2006), and path-finding models that ask how axons may optimally find their way through the nervous system (Mortimer et al., 2009).

Useful Models at the Different Levels

Many different modeling frameworks exist within motor control, including dynamical systems (Haken et al., 1985), optimal control (Todorov & Jordan, 2002), uncontrolled manifold hypothesis (Scholz & Schoner, 1999), equilibrium point control (Feldman & Latash, 2005), internal models (Wolpert & Kawato, 1998), and Bayesian statistics (Kording, 2007). Popular models in each of these frameworks can be written compactly and explain data from multiple labs and numerous experiments. Although they are each clearly useful, they also tend to have limited scope. Combining aspects of each model class may allow increasing the scope.

Models That Bridge Marr Levels

A range of other models strive to bridge across Marr levels. Such models may ask if a given algorithm is close to optimal in the sense of a given computational objective or if a given implementation actually implements an algorithm. Most models are either at one of the three Marr levels or they bridge two neighboring levels.

No Competition across Marr Levels

To obtain a full understanding of the nervous system, we believe that models are needed at each of the three levels. It is important to emphasize that models at different levels do not compete with one another. For example, the Smith model of learning (Smith et al., 2006), an algorithmic model, does not technically compete with Bayesian learning models (Kording, Tenenbaum, & Shadmehr, 2007). Smith et al.'s model asks about the algorithm by which the nervous system may solve a learning problem. Kording et al.'s model asks which problem is solved by the nervous system. Both models may lead to precise descriptions of data and contribute independently to scientists' understanding of the nervous system. Scientists want to understand both the how and the why. Even if we had a perfectly working Level 3 model we still would want to develop Level 1 and Level 2 models to make sense out of the results. At the same time, models at different levels may inform one another. Implementation knowledge may render certain algorithms impossible. Similarly computational objective models constrain potential algorithms. For example, if an algorithm is unable to solve the computational problems that need to be solved, it can not be the algorithm implemented by the nervous system.

Is Falsification Even Useful?

We want to ask if the difference of typical models in physics and motor control has implications for the meaning of falsification. Assume that the analysis of our experimental data indicates that we should reject the null hypothesis that the model predictions are identical to the measured data. We argue that this falsification may be of limited use—after all, we knew before we even started that the model was false. Moreover, the space of all possible models in motor control is huge and ruling out a single model does not significantly reduce the set of possible models. Lastly, it is arguable that falsifying a model is simply too easy. Virtually all models fit to rich data that we have seen in our careers exhibit systematic differences. For example, typical Bayesian models do not predict biases that depend on the used hand or the eye position and yet these biases are virtually always there. Ultimately, given an arbitrary amount of data, any model is either true or false but virtually all models are false. As such, finding a model to be false gives less than one bit of information and thus may have limited scientific use.[14]

Given these issues, we may argue that a good experiment should try "not to demonstrate falsity but rather to learn particular aspects of falsity" (Gelman, 2005). For example, if we find a deviation from a (Level 1) model that, say, argues that the nervous system deals optimally with noise, depending on the aspect of the falsity we might conclude that noise has a different structure than we assumed or alternatively that aspects of Levels 2 and 3 affect behavior—and we can come up with a new model. For example, we may see a reflection of the neural aspects of motor noise or of equilibrium point control in the data. We are thus of the opinion that a careful characterization of the deviation between models and experimental data is most enlightening and that the mere falsification by itself is not actually all that interesting.

Analyzing the Falseness of a Model

Models of different kinds need to be falsified[15] in somewhat different ways. We describe potential strategies of model falsification at the different Marr levels. Our discussion of such falsification has a particular focus on Bayesian models. These models have withstood a number of experimental attempts at falsification and have been falsified by other experiments.

Falsifying a Marr Level 3 (Implementation) Model

To refute statements at this level it is necessary to show that the implementation is different. This can be done in two ways: either using a physiological method that directly analyzes a property of the implementation or alternatively by testing a behavioral prediction that emerges from the model. For example, it may be argued that most motor noise arises from the muscles themselves and not the nervous system. If electrical stimulation of the muscles would produce little noise in comparison to voluntary contractions then the model would be falsified, which incidentally was the case (Jones, Hamilton, & Wolpert, 2002). However, such a model could also be refuted using behavior alone. If there was an experimental manipulation (e.g., hypnosis) that could allow people to get rid of their motor noise then muscles can not be the main source of motor noise. However, in most cases behavioral refutations are quite difficult because individual aspects of implementation are typically far removed from observable behavior. Marr Level 3 models can usually be refuted using physiological manipulations and possibly by behavioral experiments.

Falsifying a Marr level 2 (Algorithm) Model

It may be argued that the nervous system controls the forces produced by its muscles. Alternatively, it could set the equilibrium points and use this variable as the main control variable (Feldman, 1986; Feldman & Levin, 2009; Flash, 1987; Latash, 1992). One way of refuting the model that claims that the nervous system only controls forces is by introducing a position perturbation. If the nervous system controls only force, then this perturbation should not be counteracted rapidly. If the perturbation is counteracted then the pure force control model is falsified (Mussa-Ivaldi, Hogan, & Bizzi, 1985). However, an alternative way of disproving the model could have been by using electrophysiology. If we could record the descending commands in the spinal cord and find that they encode equilibrium points we would have equally disproved the pure force control hypothesis. Another example is the Smith model of two-timescale learning (Smith et al., 2006). Smith et al. argued that savings (the speedup of learning of a task that we have learned before) can be explained by the interplay of learning at different time scales. This model and similar (Marr Level 1) multi-time-scale models (Kording, Tenenbaum, et al., 2007) make the prediction that long periods of unperturbed movements should erase the savings effect. However, a recent experimental paper found this deletion of memory to not happen (Zarahn, Weston, Liang, Mazzoni, & Krakauer, 2008), thus falsifying the previously simple explanation of savings. Both behavioral and physiological methods can be used to falsify a Level 2 model.

Falsifying a Marr Level 1 (Computational Objective) Model

What is the problem that is solved by the nervous system and how well does it solve this problem? Marr Level 1 models of behavior state that the nervous system should use a good solution to the problems encountered. In no way does the objective specify if the problem should be solved by the biomechanics of the system, by neural feedback circuits or by synaptic effects. In theory, physiological methods could be used. For instance, recording all neurons and realizing that none of them encodes uncertainty would disprove Bayesian models that assume that uncertainty is a central variable.

However, this kind of falsification is very difficult. Marr Level 1 models of motor behavior can most readily be falsified by analyzing behavior.

We want to briefly discuss falsification of Bayesian models. If we have two cues that we can use for an estimation, then we should obviously rely more on the more reliable cue. For two cues that are affected by Gaussian noise and relate to the same continuous variable, Bayesian statistics makes a quantitative prediction: every cue should receive a weight that is proportional to its inverse variance. For example when we have a visual cue with uncertainty σ_V and an auditory cue with uncertainty σ_A the weight we should place on vision is:

$$W_V = \frac{1/\sigma_V^2}{1/\sigma_V^2 + 1/\sigma_A^2}. \quad (1)$$

This model can easily be falsified. We can measure the precision of vision in one experiment. We can measure the precision of audition in another experiment. We can then predict the weight subjects should place on vision using Equation 1 and then try to falsify the model by weights measured using cue conflict. Many researchers tried to falsify this model in many modalities and were generally unable to do so (Ernst & Banks, 2002; Ernst & Bulthoff, 2004; Geisler & Kersten, 2002; Kersten & Yuille, 2003; Rosas, Wagemans, Ernst, & Wichmann, 2005; van Ee, Adams, & Mamassian, 2003). However, clearly this model must be incomplete because it assumes that the two cues are always related to one another. Indeed, whenever two cues are sufficiently different the model breaks down and subjects perceive two independent perceptual events (Gepshtein, Burge, Ernst, & Banks, 2005; Jack & Thurlow, 1973; Kording, Beierholm, Ma, Quartz, Tenenbaum, et al., 2007; Munhall, Gribble, Sacco, & Ward, 1996; Roach, Heron, & McGraw, 2006; Sato, Toyoizumi, & Aihara, 2007; Slutsky & Recanzone, 2001; Wei & Kording, 2008). So clearly the simple Bayesian model of cue combination is falsified and other principles, such as causality, may become relevant.

Depending on the Marr level a model is formulated on, it is possible to falsify the model using different classes of experiments. However, what we learn from falsification depends on the kind of model considered. Because models in motor control appear to be either "false" or not compact falsification strategies need to be different.

Implications for Experiments and Models

If falseness per se is not that important—as argued previously—but it is the nature of the difference that matters, then experiments should aim at characterizing this difference. At present there is intense pressure on experimentalists to perform hypothesis-driven research. Virtually all granting agencies and journals demand that researchers should try to disprove a given hypothesis. However, there is much less drive to characterize the deviations. It may be argued that giving more credit to scientists who carefully characterize deviations and share the resulting data could accelerate progress in motor control.

It is often argued that published figures in papers should be sufficient but there are obvious problems with that approach. Specifically, figures in published papers are generally chosen so that they highlight those effects that the researchers found interesting. These phenomena may not be the ideal phenomena for testing a model. Therefore, we argue that the best way of testing models is by verifying how well they explain the raw experimental data. This is only practically possible by the adoption of databases by the field of motor control. Databases have been highly useful in various communities (e.g., RCSB PDB, http://www.rcsb.org/). Such databases help ensure that data is reproducible on top of providing a way of quantitatively assessing the amount of data explained by each model.

If it is the objective of models to be good approximations in the sense that they explain a lot of data variance, then models should apply to as much data as possible. However, most models in motor control are applied to a single data set only. With this restricted information it is difficult to assess how much variance is explained by a given model. In order to facilitate this process, models should be shared across labs. It may be argued that it would be helpful to make it mandatory to publish computer programs that implement models along with published modeling studies. This would facilitate model comparison and could thus also accelerate progress in motor control.

In this article we argue that the difference between typical models in physics and typical models in motor control may be fundamental and that this difference may have implications for the falsification of models. We propose that model development should strive to explain experimental variance and thus to be as good an approximation as possible. We also propose that the difference between models has implications for experiments, which should aim to analyze the nature of falseness of models.

ACKNOWLEDGMENTS

Hugo L. Fernandes is supported by the PhD Program in Computational Biology of the Instituto Gulbenkian de Ciência, Portugal (sponsored by Fundação Calouste Gulbenkian, Siemens SA, and Fundação para a Ciencia e Tecnologia; SFRH/BD/33525/2008). The authors thank the NIH (R01NS057814, R01NS063399) and the Chicago Community Trust for financial support. They also want to thank Bruno Martins for helpful discussions and an anonymous referee for inspiration. Last, the authors want to thank the 2009 NCM satellite workshop on falsifiability in motor control organized by Dagmar Sternad and Robert Ajemian without which we certainly would not have written the article. They want to thank John Krakauer, Opher Donchin, and Mark Latash for outstandingly fruitful discussions.

NOTES

1. We need to emphasize that we are scientists studying motor control and have no formal training in the philosophy or sociology of science. As such, we want to make clear that this

article is work in progress and only reflects our present thinking behind the practical studies we perform in the area of motor control. We also want to stress that there are a good numbers of attempts at comparing physics and biology (e.g., Schrodinger, 1946) and that we only use a superficial comparison as a starting point of our discussion.

2. It is hard to define what a model is. In the context of this article we simply consider it an approximation to reality.

3. Compactness is a concept that is difficult to define. Intuitively compactness is the number of book pages we would need to fully write out the model, including all parameters. A more mathematical specification would be the length of the shortest Turing program to implement the model, a concept equivalent to Kolmogorov complexity (Li & Vitanyi, 1997).

4. In this article, we give a lot of reasons for our believe in this hypothesis, but obviously we could be wrong. Our hypothesis could be falsified by someone coming up with a compact model that precisely explains the bulk of data on motor control.

5. We focus on data variance in this paper because no model explains all the data and the world is probabilistic. Although we talk about data variance we acknowledge the structure of noise and ask which model makes the data most likely.

6. The question if a model should describe some data very well or a wide range of data acceptably is hard to answer and outside of the scope of this article.

7. In fact, one of the authors of this article, Konrad Kording, spent most of his PhD thesis building biologically realistic models of cortical function and thus models that fall squarely into this class.

8. We want to emphasize that although "true" models need to be bottom-up, there are "false" bottom-up models. For example, a bottom-up model may assume random synaptic weights—clearly not meant to be true—but which may be a sufficient assumption for say modeling network dynamics.

9. One possibility of reducing the number of free parameters is to build hybrid models in which parameters are determined by implicit top-down models. Such models have some properties of "true" models and some properties of "false" models and may be feasible. However, as this approach is not formalized yet in a clear way, we do not discuss it in detail in the present article.

10. In fact, it may be argued that a central difference between physics and neuroscience is the nature and usefulness of approximations. For example, many approximate models in physics approach reality arbitrarily well in certain limits. For example, ideal gas equations become arbitrarily precise for low density and high temperature gases. Analogous approximations in neuroscience (e.g., mean field or random connectivity) do not seem to have such properties. This may render models in the motor control domain trivially falsifiable.

11. For answering certain scientific questions, certain aspects of the nervous system may not be relevant. For example, the exact pattern of neural connections may not be relevant if we want to predict the patterns of epilepsy (Traub & Wong, 1982). Those models may well be compact—however, the underlying approximations make these models fall into the "false" category.

12. Incidentally, this classification does not apply to physics. Physical systems do not generally have a purpose but certainly many systems in biology are best understood by assigning purpose to them. Marr Level 1 is explicitly concerned with computational purposes.

13. The Marr levels are a model of models. As such, we do not know how complete the Marr levels are and there may well be models to which the Marr levels cannot be gainfully applied.

14. Although experimental falsification per se is not very interesting, we want to stress that it is absolutely vital for models to be falsifiable (Popper, 1959). If a model only makes nontestable hypotheses it does not help scientific progress. From our perspective it is of primary importance for a model to predict data.

15. In the following we use the word *falsify* to refer to the process of analyzing how a model is false.

REFERENCES

Borges, J. L. (1999). *Borges: Collected fictions*. New York: Penguin.

Carandini, M., Demb, J. B., Mante, V., Tolhurst, D. J., Dan, Y., Olshausen, B. A., et al. (2005). Do we know what the early visual system does? *Journal of Neuroscience*, *25*, 10577–10597.

Chen, B. L., Hall, D. H., & Chklovskii, D. B. (2006). Wiring optimization can relate neuronal structure and function. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 4723–4728.

Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. Cambridge, MA: MIT Press.

Ernst, M. O., & Banks, M. S. (2002, January 24). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429–433.

Ernst, M. O., & Bulthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*, 162–169.

Feldman, A. G. (1986). Once more on the equilibrium-point hypothesis (lambda model) for motor control. *Journal of Motor Behavior*, *18*, 17–54.

Feldman, A. G., & Latash, M. L. (2005). Testing hypotheses and the advancement of science: Recent attempts to falsify the equilibrium point hypothesis. *Experimental Brain Research*, *161*, 91–103.

Feldman, A. G., & Levin, M. F. (2009). The equilibrium-point hypothesis—Past, present and future. *Advances in Experimental Medicine and Biololgy*, *629*, 699–726.

Feynman, R., Leighton, R., & Sands, M. (1963). *The Feynman lectures on physics, vol. 1*. Reading, MA: Addison Wesley.

Flash, T. (1987). The control of hand equilibrium trajectories in multi-joint arm movements. *Biological Cybernetics*, *57*, 257–274.

Fusi, S., Drew, P. J., & Abbott, L. F. (2005). Cascade models of synaptically stored memories. *Neuron*, *45*, 599–611.

Geisler, W. S., & Kersten, D. (2002). Illusions, perception and Bayes. *Nature Neuroscience*, *5*, 508–510.

Gelman, A. (2005, September 2). *One more time on Bayes, Popper, and Kuhn*. Retrieved from http://www.stat.columbia.edu/cook/movabletype/archives/2005/05/one_more_time_e_2.html

Georgopoulos, A. P., Schwartz, A. B., & Kettner, R. E. (1986, September 26). Neuronal population coding of movement direction. *Science*, *233*, 1416–1419.

Gepshtein, S., Burge, J., Ernst, M. O., & Banks, M. S. (2005). The combination of vision and touch depends on spatial proximity. *Journal of Vision*, *5*, 1013–1023.

Gielen, C. C., & Houk, J. C. (1987). A model of the motor servo: Incorporating nonlinear spindle receptor and muscle mechanical properties. *Biological Cybernetics*, *57*, 217–231.

Haken, H., Kelso, J. A., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics*, *51*, 347–356.

Jack, C. E., & Thurlow, W. R. (1973). Effects of degree of visual association and angle of displacement on the "ventriloquism" effect. *Perceptual Motor Skills*, *37*, 967–979.

Jones, K. E., Hamilton, A. F., & Wolpert, D. M. (2002). Sources of signal-dependent noise during isometric force production. *Journal of Neurophysiology*, *88*(3), 1533–1544.

Kandel, E., Schwartz, J., & Jessell, T. (2000). *Principles of neural science*. New York: McGraw-Hill/Appleton & Lange.

Kersten, D., & Yuille, A. (2003). Bayesian models of object perception. *Current Opinion in Neurobiology*, *13*, 150–158.

Kording, K. P. (2007, October 26). Decision theory: What "should" the nervous system do? *Science*, *318*, 606–610.

Kording, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS ONE*, *2*(9), e943.

Kording, K. P., Tenenbaum, J. B., & Shadmehr, R. (2007). The dynamics of memory as a consequence of optimal adaptation to a changing body. *Nature Neuroscience*, *10*, 779–786.

Kuhn, T. (1970). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Kurt, H., Maxwell, S., & Halbert, W. (1990). *Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks*. San Diego, CA: Department of Economics, University of California, San Diego.

Latash, M. L. (1992). Independent control of joint stiffness in the framework of the equilibrium: Point hypothesis. *Biological Cybernetics*, *67*, 377–384.

Li, M., & Vitanyi, P. (1997). *An introduction to Kolmogorov complexity and its applications*. New York: Springer-Verlag.

Maass, W., Natschläger, T., & Markram, H. (2002). Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation*, *14*, 2531–2560.

Markram, H. (2006). The blue brain project. *Nature Reviews Neuroscience*, *7*, 153–160.

Marr, D. (1982). *Vision: A computational approach*. San Francisco: Freeman.

Morrow, M. M., Jordan, L. R., & Miller, L. E. (2007). Direct comparison of the task-dependent discharge of M1 in hand space and muscle space. *Journal of Neurophysiology*, *97*, 1786–1798.

Mortimer, D., Feldner, J., Vaughan, T., Vetter, I., Pujic, Z., Rosoff, W. J., et al. (2009). Bayesian model predicts the response of axons to molecular gradients. *Proceedings of the National Academy of Sciences of United States of America*, *106*, 10296–10301.

Munhall, K. G., Gribble, P., Sacco, L., & Ward, M. (1996). Temporal constraints on the McGurk effect. *Perception & Psychophysics*, *58*, 351–362.

Mussa-Ivaldi, F. A., Hogan, N., & Bizzi, E. (1985). Neural, mechanical, and geometric factors subserving arm posture in humans. *Journal of Neuroscience*, *5*(10), 2732–2743.

Popper, K. (1959). *The logic of scientific discovery*. London: Hutchinson.

Prochazka, A., & Gorassini, M. (1998). Models of ensemble firing of muscle spindle afferents recorded during normal locomotion in cats. *Journal of Physiology*, *507*, 277–291.

Roach, N. W., Heron, J., & McGraw, P. V. (2006). Resolving multisensory conflict: A strategy for balancing the costs and benefits of audio–visual integration. *Proceedings of the Royal Society B: Biological Sciences*, *273*, 2159–2168.

Rosas, P., Wagemans, J., Ernst, M. O., & Wichmann, F. A. (2005). Texture and haptic cues in slant discrimination: Reliability-based cue weighting without statistically optimal cue combination. *Journal of the Optical Society of America A*, *22*, 801–809.

Sato, Y., Toyoizumi, T., & Aihara, K. (2007). Bayesian inference explains perception of unity and ventriloquism aftereffect: Identification of common sources of audiovisual stimuli. *Neural Computation*, *19*, 3335–3355.

Scholz, J. P., & Schoner, G. (1999). The uncontrolled manifold concept: Identifying control variables for a functional task. *Experimental Brain Research*, *126*, 289–306.

Schrodinger, E. (1946). *What is life?* Cambridge: Cambridge University Press.

Slutsky, D. A., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *Neuroreport*, *12*, 7–10.

Smith, M. A., Ghazizadeh, A., & Shadmehr, R. (2006). Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biology*, *4*(6), e179.

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, *5*, 1226–1235.

Traub, R. D., & Wong, R. K. (1982, May 14). Cellular mechanism of neuronal synchronization in epilepsy. *Science*, *216*, 745–747.

van Ee, R., Adams, W. J., & Mamassian, P. (2003). Bayesian modeling of cue interaction: Bistability in stereoscopic slant perception. *Journal of the Optical Society of America A*, *20*, 1398–1406.

Wei, K., & Kording, K. P. (2008). Relevance of error: What drives motor adaptation? *Journal of Neurophysiology*, *101*, 655–664.

Wolpert, D., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, *11*, 1317–1329.

Zarahn, E., Weston, G. D., Liang, J., Mazzoni, P., & Krakauer, J. W. (2008). Explaining savings for visuomotor adaptation: Linear time-invariant state-space models are not sufficient. *Journal of Neurophysiology*, *100*, 2537–2548.

Submitted September 23, 2009
Accepted June 30, 2010