Overview

Bayesian approaches to sensory integration for motor control

Max Berniker and Konrad Kording*

The processing of sensory information is fundamental to the basic operation of the nervous system. Our nervous system uses this sensory information to gain knowledge of our bodies and the world around us. This knowledge is of great importance as it provides the coherent and accurate information necessary for successful motor control. Yet, all this knowledge is of an uncertain nature because we obtain information only through our noisy sensors. We are thus faced with the problem of integrating many uncertain pieces of information into estimates of the properties of our bodies and the surrounding world. Bayesian approaches to estimation formalize the problem of how this uncertain information should be integrated. Utilizing this approach, many studies make predictions that faithfully predict human sensorimotor behavior. © 2011 John Wiley & Sons, Ltd. *WIREs Cogn Sci* 2011 2 419–428 DOI: 10.1002/wcs.125

INTRODUCTION

A central objective of the nervous system is to sense the state of the world around us and to affect this state such that it is more favorable to us. To achieve this, our senses provide us with information about our bodies and our immediate environment. All this noisy information must be integrated into a cohesive whole before we can act rationally. Because of its importance, the integration of sensory information has been studied frequently both in the realm of *perception* and in the realm of *sensorimotor integration*. Sensorimotor integration in particular has been widely studied under the banner of *motor control*. Indeed, it is through our motor actions that we can most readily bring about changes in the state of the world.

Motor control refers to nervous system's role in planning, executing, and stabilizing the movements of the body. Studies of motor control generally ask how the nervous system uses sensory information to achieve its goals. This requires knowledge about our motor apparatus, the world around us, and how the two interact. The *motor control problem* is then how to use our sensory information to best elicit a desired outcome. A crucial first step for motor control is therefore to *integrate sensory information* reliably and accurately.

*Correspondence to: kk@northwestern.edu

Northwestern University, Rehabilitation Institute of Chicago, Chicago, IL, USA

DOI: 10.1002/wcs.125

Below we describe how the problem of integrating sensory information is central to the motor control problem and how its mathematical description necessitates a statistical framework for examination: *Bayesian statistics*. We start by describing the general motor control problem and the need for a statistical framework. This will set the stage for the various experimental approaches that have been taken to examine if subjects perform similar statistical inferences under a wide variety of circumstances. Once these inferences have been made, computational methods, covered in this review, can be employed to find appropriate motor commands. In this article, however, we focus on how Bayesian statistics can be used to formalize the process of estimation, the necessary first step before efficient motor commands can be produced.

THE MOTOR CONTROL PROBLEM

Humans and other animals typically produce movements in a stereotyped fashion. This implies that the nervous system has some notion that certain commands and movements are better than others—even when the commands and movements in question produce the same outcome. Within the framework of *decision theory* and *optimal control*, the notion that certain commands are better than others is formalized by a *cost function*: a function that quantitatively evaluates the merit of a command and its resulting outcome. Given this very general, although somewhat abstract description, we can mathematically formulate the problem of motor control as a minimization. For illustrative purposes, we first present this in an overly simplified manner to motivate the general approach. Then we add the necessary further details to introduce the statistical nature of the problem. First, though, we can state the problem in simple terms: which command u^* results in the minimum cost?

$$u^* = \arg\min_u \{\text{cost}(\text{outcome}(u), u)\} \tag{1}$$

For instance, suppose we are throwing darts. The 'cost' associated with the dart's position on the board should decrease with increasing scores; that is, we achieve the lowest cost when we get the highest score. Therefore, the best command would be that which delivers the dart to the triple 20 (60 points, better than the bull's eye) and minimizes the cost (and maximizes our score).

The above example assumes that we can deterministically choose a command and be sure of its outcome. Yet, most motor control problems are *stochastic* in nature, and motor outcomes are affected by *uncertainty*. Under more realistic conditions our minimization problem is no longer equivalent to Eq. (1), but must take into account the statistics of the task. To continue the example, if we aim at the center of the dartboard we are not certain to hit the bull's eye with the dart (even if we are a seasoned champion). If we continue to aim at the center and throw many times we end up with a distribution of dart positions. This distribution, $p(\text{outcome} = x|u)$ characterizes the likelihood of each outcome, the dart landing at position x given that we chose the command u. This distribution helps to define our new minimization problem: what command minimizes the likely cost?

$$u^* = \arg\min_u \left\{ \sum_{\text{possible outcomes}} \text{cost}(\text{outcome}(u), u) p(\text{outcome}|u) \right\}. \tag{2}$$

The optimal command thus defined minimizes the so-called *expected cost*, the sum of possible costs weighted by their corresponding probabilities. In the example of playing darts, the best aiming point is a point where we receive high scores on average, even if we make large mistakes. This best aiming point will change from player to player depending on their likelihood of landing on any particular target. Indeed, both amateur and world-class players are known to adopt a strategy that is well predicted by this approach.

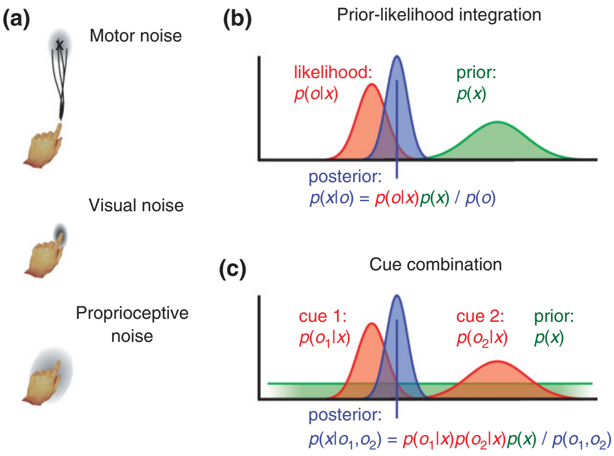

FIGURE 1 | (a) Possible sources of noise in the motor system. Motor commands possess noise and result in uncertain movements. Visual and proprioceptive senses too, contain inherent uncertainties. (b) Bayesian integration of a prior and likelihood. The prior, denoted with the green curve, represents the probability of a state, x. The likelihood, denoted with the red curve, represents the probability of observing the data, o, given x. The posterior is the probability that x is the state, given our observation, o. (c) Bayes' rule applied to cue combinations is the mathematical analog, only instead of using a prior and likelihood, we integrate two likelihoods. Here, for simplicity we assume the two observations are conditionally independent given the state, and a flat prior distribution over x.

While playing darts, our uncertainty in the dart's location arises largely from motor noise (Figure 1(a)). However, there are many sensory sources of uncertainty as well. For example, our visual system is noisy and our sense of the location of the dartboard relative to our body is uncertain (Figure 1(a)). Moreover, our proprioceptive system is noisy as well; the orientation of our hand and arm as we release the dart are uncertain (Figure 1(a)). These and many other sources of uncertainty[1] combine to produce variability in the motor outcome, given our command decision.

Continuing our example, we can examine more sophisticated and realistic motor control problems. For instance, a more sensible description of dart throwing recognizes that the task is dynamic. The dart's final position on the board depends on the ballistic trajectory it takes once it has been released from our hand. The motion of the dart up until the moment we release it is dictated by the inertial mechanics of our limb and the force generating properties of our muscles. Clearly, our description of the dart-throwing problem can take on greater and greater levels of detail and physical accuracy. Nevertheless, the motor control problem is still to solve for a command, u, although one that varies in time due to the dynamic nature of the problem. The minimization to be performed at each instant is a sum over both the statistics of the possible outcomes at the current time, and the statistics of future possible outcomes that unfold as a result of the current choice of u.

For dynamical tasks such as these, we need to know the *state* of the system and the world to compute the optimal motor command. For the dart-throwing task the state could be the orientation and velocity of our hand. To measure the state we use sensory feedback, but just as with motor outcomes, any measurement of the state is corrupted by noise. Therefore motor controllers must cope with the further uncertainty of the state as well as how this influences motor outcomes.

Despite the complexity of the above examples, we can formulate the dynamical motor control problem in an expanded version of Eq. (2). We must solve a statistical problem concerning likely motor outcomes conditioned on our choice of commands. The difficulty with this problem, and motor control in general, is that computing the probability of an outcome given a motor command is difficult. Doing so demands the specification of many variables that are uncertain and subject to their own statistics. This includes information about our body, the world, and how they interact. Integrating all this uncertain information requires a statistical approach: *Bayesian statistics*. Bayesian integration is the mathematical framework that calculates how uncertain information from multiple sources can be optimally combined. It results in a coherent and maximally accurate estimate of a set of observations. Using Bayes' rule we can integrate multiple measurements, or multiple pieces of distinct information about a variable, into a new probability. We can also update the probability of a variable over time as we continue to gain new information about it. Finally, we can also use our observations and Bayesian integration to update our beliefs about the structure of the world; that is, what processes are responsible for shaping our observations.

Given the fundamental implications for how estimated properties of our bodies and the world impact our ability to make optimal motor decisions, exactly how human subjects integrate their noisy and uncertain sensory information is of great interest. The issue of computing an optimal control signal based on such information can also be examined in a Bayesian framework. A number of recent psychophysics studies analyze how people and animals integrate multiple sources of uncertain information to make sensorimotor decisions. Below we discuss some of these findings. We focus on circumstances that are sufficiently simple that coherent statistical models can be used to predict human behavior.

BAYESIAN INTEGRATION

Combining Prior Knowledge with New Evidence

Combining uncertain information to produce a coherent and accurate estimate of our body and the world is a central problem in motor control. As an example, consider the task of descending a staircase. Based on our familiarity with walking downstairs, we have strong assumptions for things like the distance between steps, their height, and their general shape. These assumptions form a prior over stairs, a belief in their typical properties. Often these priors are strong enough that we feel comfortable taking stairs without even observing them, as when we descend stairs without looking at our feet, or in the dark. Normally though, we will first observe how far down we will need to step. However, vision does not provide perfect information. The visual system provides us with a likelihood of the step's height. This likelihood is the probability of having a particular sensory observation for each possible stair height. Bayes' rule defines how to combine the prior and the likelihood to make an optimal estimate of the step's height.

Bayes' rule states that the probability of the step's height being value x, given our observation, o, is the product of the prior probability of the stair height and the likelihood, normalized by the probability of the observation. Mathematically, this is expressed as:

$$p(x|o) = p(x)p(o|x)/p(o). \quad (3)$$

The distribution produced by Eq. (3) is known as the posterior probability (this is shown graphically in Figure 1(b)). We can also interpret Bayes' rule as the 'optimal' means of combining a prior and a likelihood, as it results in a distribution with minimal uncertainty.

Several studies, using many sensory modalities, have shown that when subjects combine preceding knowledge with new information their behavior reflects the integration of a prior and likelihood in a manner prescribed by Bayes' rule. A typical study[2] will have people indicate their estimate of a target's location through a motor task. In each trial, the target's location is drawn from a Gaussian probability distribution (the prior). The distribution can be fixed, or vary across subjects as an experimental condition. Noisy feedback of the target's location is provided (the likelihood). This distribution too, can be used as an experimental condition. Bayesian statistics predicts how subjects should combine the likelihood and the prior. These predictions are then compared to human performance.

These paradigms have been applied to a wide range of topics spanning sensorimotor integration, force estimations, timing estimations, speed estimations, stance regulation, the interpretation of visual scenes, and even cognitive estimates.[3–11] Together these studies demonstrate that people are adept at combining prior knowledge with new evidence in a manner predicted by Bayesian statistics.

Combining Multiple Pieces of Information

In many cases it is not prior knowledge that is combined with the likelihood, but rather two different pieces of information that are combined. For example, we may see and feel an object at the same time.[12] We can then use what we saw and what we felt to estimate a common property of the object, for instance, its size or its texture. This type of task is what is commonly referred to as *cue combination*; two or more sensory cues are combined to form a common, or joint, estimate. Only by accurately combining these sensory cues can we optimally complete the task. Just as before, Bayesian statistics prescribes how we should combine the likelihoods to compute an optimal estimate from the posterior distribution (Figure 1(c)).

Recent studies have found that when combining information this way, people are also similar to optimal. As an example study, we consider how people combine visual and auditory information to estimate the position of a target. First the precision of visual and auditory perceptions are separately measured for each subject.[13] This is done to characterize the subject's likelihood for the two sensory modalities. Then the precision and accuracy of perception is measured when subjects use both senses. Their performance in these cue-combination trials can be predicted using the rules of Bayesian integration, further evidencing people's ability to optimally cope with uncertain information.

In addition to auditory and visual cues, the combination of cues from different modalities has been analyzed in a good number of studies.[7,12,14–27] In typical cases cues are combined by the subjects in a fashion that is close to the optimum prescribed by Bayesian statistics.

Credit Assignment

The rules of Bayesian statistics reviewed above describe how we can bring new information to bear upon our beliefs. As in the examples above, sometimes we are concerned only with estimating a single property. However, under some cases we are concerned with estimating many properties. For instance, suppose we hold two blocks in our hand, one sitting atop the other, and we want to estimate their individual weights. Our observation of their combined weight upon our hand is indicative of their individual weights, namely the magnitude of their sum. However, this information is not sufficient to accurately establish their individual weights. To estimate them, we need to solve a credit assignment problem; that is, how does each property (individual weight) contribute to our observation (overall weight). Bayesian statistics also prescribes an optimal solution to this problem.

Our observation of their combined weight forms a likelihood, a probability of having a particular sensory observation of their combined weight, for each value of their individual weights. By combining this with a prior over the object's weight, perhaps based on their size,[28] we can compute a posterior distribution of their individual weights. Again, this distribution will provide an optimal estimate of their individual weights, simultaneously solving the credit assignment problem: how much does each block contribute to the total weight I feel? Recent research[29] addresses this kind of problem by integrating the information over time.

BAYESIAN INTEGRATION OVER TIME

In our discussion above we were concerned with estimating what we assume are static properties such as the weight of a block or the location of an auditory cue. However, the world is dynamic, and as such its properties and our perception of them are continually changing. Thus we constantly need to integrate this information with our current beliefs to inform new estimates of the world if we are to behave optimally. This implies that Bayesian integration should take place in a continuous manner.

This approach is taken extensively in modern applications of control theory through the use of a Kalman filter. Kalman filtering is a procedure for using a model and our observations to continuously update our beliefs (Figure 2(a)). This technique, though not usually described as such, is equivalent to an application of Bayesian integration. At each instant the Kalman filter combines the model's estimate of the world's state (the prior) with a measured observation (the likelihood) to update a prediction of the world's current state, represented by a posterior (Figure 2(b)). At any point of time, the posterior from the past is combined with the model's dynamics and the motor command to define the prior for the future. This formalism is well developed and used in a wide range of applications from aeronautics to humanoid robotics.[30,31] Indeed, even the motor control problems of two applications as disparate as controlling a jet and controlling our bodies, share many computational analogies. In both cases continuously incoming

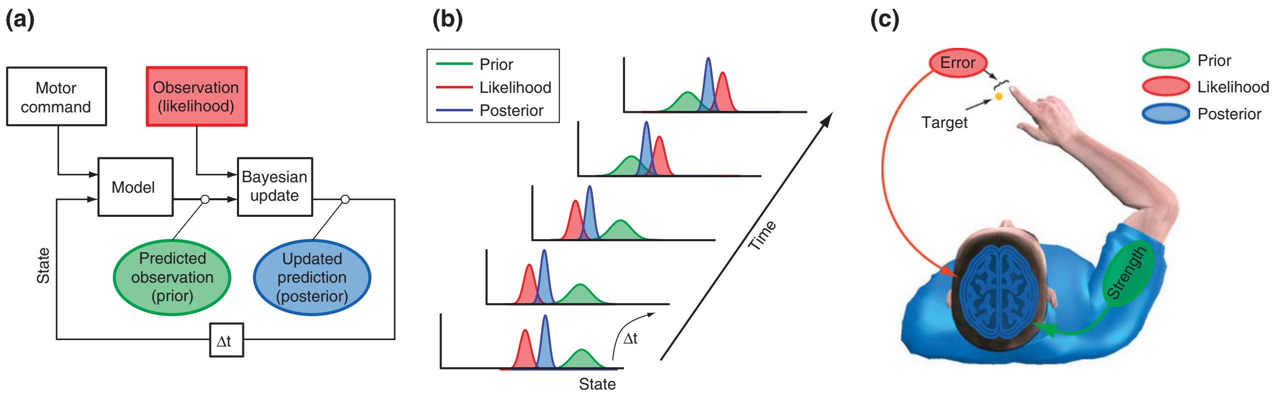

FIGURE 2 | (a) Typical procedure for optimally estimating the world's state in modern control theory. A model of the world, combined with a motor command is used to estimate a predicted state (and its observation), through a prior distribution (labeled in green). Observations of the world dictate the likelihood of a particular observation given the current world state (labeled in red). A Kalman filter is used to make a Bayesian update of our belief in the world's current state (labeled in blue). (b) This process of Bayesian inference repeats itself at each time step, using the posterior from one time step, as the prior for the following time step. (c) Motor adaptation can be framed as an analogous update procedure. Our prior belief in muscle properties (e.g., muscle strength, labeled in green) is integrated with our observed motor errors (labeled in red) to update estimates in our muscle properties.

information needs to be assessed to move precisely (*albeit* on different time scales). Only recently has the methodology of Kalman filtering been applied to make quantitative predictions of human movement behavior.

As an example of this problem we consider a recent theory of motor adaptation.[29] The properties of our bodies change continuously throughout our lives. For instance, our muscles can fatigue over the course of minutes, and grow or shrink in strength over the course of weeks. This means that we must estimate the strength of our muscles if we are to move precisely. Errors in the estimated strength of our muscles will translate into movement errors. We can use this error to obtain a likelihood function characterizing how strong the muscles are. According to Bayesian statistics we need to combine newly obtained information, our motor error, with our prior belief.[32,33] We can thus infer a new and improved estimate of our muscles (Figure 2(c)).

It has been found that motor adaptation over time can be well understood from the predictions of Kalman filtering. For instance, evidence suggests how people estimate the position of their hand,[34] adapt to robot-rendered force fields,[29] and even balance a pole upon their hand[35] can all be explained using this strategy. Taken together, these studies highlight that when people integrate information over time, they seem to do so in a fashion that is consistent with the optimal Bayesian solution. These findings have important consequences for optimal control, where computing the optimal command relies on accurate estimates of the world's state.

BAYESIAN INFERENCE OF STRUCTURE

Above we discussed how people appear to combine two cues into a joint estimate using Bayes' rule. This is a necessary initial step in any sensorimotor task. The analyses of these tasks implicitly assume that subjects are certain that both cues have a common source: if we hear a tone while seeing a flash of light, we assume that both of these cues originate at the same location and have a common cause. This idea has been very successful at describing human behavior in a variety of cue-combination experiments. Yet intuitively, we recognize that before we can form estimates from our perceptions, we must infer the underlying structure of our sensory stimuli, only then can we make informed motor decisions.

Based on common experience and experimental studies we know that if two cues are very dissimilar, our perception of a common source breaks down. If we see a flash far to our right while perceiving a tone far to our left, we may perceive two independent events. We estimate if two events have a common cause or if they just randomly co-occur; that is, we infer the structure of events. One result of this inference is the so-called ventriloquist effect.[36–39] A commonly experienced example is found while watching television: we experience the illusion that the characters voices emanate from their mouths (and not your television speakers). However, the illusion quickly breaks down if the timing of the program's sound track is off. This implies that if two events are perceived as having a common cause they should influence the perception of each other. On the other hand, if they are perceived as not having a common cause the perception of each should be independent of the other.[40,41]

Traditional Bayesian analyses of psychophysical behavior used tasks where the experimental cues were close to coincident. For example, the positions of visual and auditory stimuli were usually closely located.[13] Examining a richer statistical problem, new studies have tested subject performance in situations where two cues are dissimilar from one another. A typical study will present a visual stimulus at a random position within the visual field and simultaneously present an auditory stimulus at a different random position. Subsequently subjects are asked where they perceive the stimuli. These studies have found that when two stimuli are near coincident, people tend to infer a common cause and use each cue to guide the estimation of the other (just as in the traditional cue-combination experiments). With increasing disparity between the cues subjects' belief in a common cause decreases (Figure 3). Traditional Bayesian models, relying on an assumption of a common cause, predict that the value of one observed cue should be linearly influenced by the value of the other cue (when assuming Gaussian likelihoods). In sharp contrast to these predictions, new experiments have found nonlinear cue interactions. However, these interactions are clearly explained if we assume that on each trial the subjects infer a common or distinct cause for both cues.[40,41]

These effects are found in a wide range of experimental situations. They have been found in position estimation using visual and auditory stimuli,[40] in the estimation of depth from several cues,[20] in visuomotor adaptation,[42] and also in experiments where subjects have to estimate the number of events, for example, of visual, tactile, and auditory stimuli.[43] The evidence suggests that not only do subjects integrate priors and likelihoods but also the causal structure underlying a set of sensory stimuli.

The statistical problem that subjects solve in cue combination implicitly involves an inference about the causal structure of the stimuli. In these studies people are uncertain about the identity and number of relevant variables, for example, did the flash of light causally give rise to the auditory tone, did the tone causally give rise to the flash of light, was there an unknown event that caused them both, or were they

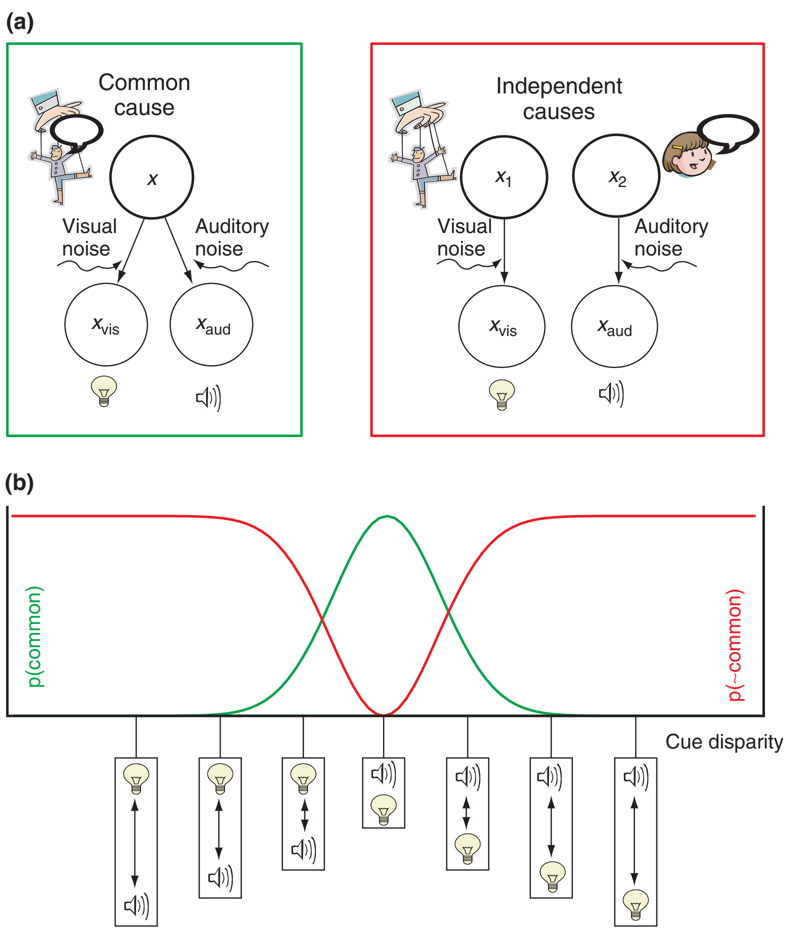

FIGURE 3 | (a) Two structural beliefs of the world. If two cues are adequately coincident, subject's perceive them as having a common cause (the green box), a phenomenon typified through ventriloquism. If the cues are disparate in time or space, subjects perceive them as having independent causes (the red box). (b) Subject's belief of a common cause. As the spatial disparity of two cues, a light flash and a tone, is experimentally controlled the belief in a common cause can be manipulated.

simply coincidental? The problem faced by the nervous system is thus similar to cognitive problems that occur in the context of causal induction.[44–51] Many experiments demonstrate that people interpret events in terms of cause and effect. The results presented here show that sensorimotor integration exhibits some of the same factors found in human cognition; only after interpreting two cues as having the same cause would they be integrated to estimate a single location. Carefully studying and analyzing seemingly simple problems such as cue combination may provide a fascinating way of studying the human cognitive system in a quantitative fashion.

LOOKING AHEAD: MAKING OPTIMAL DECISIONS

We have surveyed a great amount of evidence on how people gauge, judge, and generally make sense of, the world. We have shown that human subjects estimate the structure of the world, using models to predict how its properties evolve over time, and combine multiple sensory cues or their prior experience, to form a best estimate. These are valuable processes for our accurate perception of the world. Furthermore, as we have described, these estimations are also critical for motor control. Conceptually, however, this is only the first step in motor control. We still need to choose a motor command as shown in Eq. (2). Once we have estimated the state of the world, we need to compute the distribution over possible outcomes of any given action. Only then can we choose the command that will minimize the expected cost. This formalism of choosing optimal commands by combining the statistics of our actions and the resulting costs falls under the field of *decision theory* and *optimal control*. In all but the simplest cases, computing these optimal commands is very difficult. Despite this hindrance, many researchers have examined simple motor tasks to compare people's motor decisions with optimal decisions.

Part of the difficulty in examining these problems is in knowing beforehand what people try to optimize. Therefore, researchers have designed experiments where the cost of the task is relatively explicit such as the dartboard example. In a set of reaching studies not unlike throwing a dart, it was observed that people are remarkably close to making the optimal choices prescribed by the decision theory.[52–54] In other experiments the objective is rather implicit in the task description[55–57] and force producing tasks.[58] This further demonstrates people's abilities to integrate statistical information in a Bayes' optimal manner not only just for estimation but also for control.

DISCUSSION

The Bayesian studies we have reviewed predict the results from many experimental studies on sensory integration, cue combination, and motor control. In theory, Bayesian statistics make predictions of optimal behavior in any situation. However, there may be limitations; human subjects do not optimally speculate on, for example, the stock market.[59] Despite this, in many circumstances our motor performance in uncertain conditions appears superior to our decision making under similar conditions.[60] It may be argued that the low-level behaviors we have surveyed here, are of the kind that were evolutionarily important, or simply easy enough, that humans are good at it. Bayesian and other *normative approaches* (see below) are primarily of predictive value when we have reasons to assume that behavior needs to be good.

In large part, the studies that we have presented attempted to estimate the parameter values necessary for our motor tasks. This process often falls under the banner of motor adaptation. In addition, we have discussed how Bayesian analysis can be used to infer the causal structure of our motor tasks. These processes are often thought to be distinct from that of motor learning, which takes place on a longer time scale and requires more intense training. However, this distinction need not mean that Bayesian ideas are not applicable. Future analyses of motor learning could incorporate sophisticated Bayesian models that infer both structure and model parameters for a task. Such approaches would, by their very nature, require more evidence to converge on solutions, and be consistent with the longer time scales of motor learning.

Most Bayesian models discussed above assumed that noise sources are Gaussian and that the interactions between the subject and the world are linear. A set of recent approaches in Bayesian statistics allows the application of powerful frameworks well beyond these simple assumptions. For instance, the structure inference problem is a simple case of non-Gaussian probability distributions. The techniques that may be applied to such problems range from linearization techniques such as extended Kalman filters that represent probability distributions as Gaussians, to variational techniques that represent them as arbitrary functions, to particle filtering methods that can represent any probability distribution.[61,62] It may be hoped that more realistic models of the motor system will benefit from further advances in machine learning.

The evidence reviewed for Bayesian methods has been based on psychophysical experiments where human behavior is measured and the underlying neural processes are unobserved. The Bayesian models proposed are often interpreted as archetypes for the neural computations responsible for the motor behaviors. Recent studies though are looking for evidence of these computations at the neuronal level. Purely computational studies now show how neurons might easily solve some of the needed computations.[63,64] Other studies are looking for direct electrophysiological evidence of multimodal cue combination in the context of uncertainty.[65] Further modeling studies ask about the distributed nature of computations that might underlie the computation with uncertainty. Future studies might focus on imaging or electrophysiological studies to find more evidence for the neural substrates of these computations.

There are two broad approaches to modeling the behavior of the nervous system. In one approach, researchers start by making assumptions about the problem which the nervous system is solving. For instance, they may assume accuracy and efficiency are the goals when trying to reach a target.[66,67] Then computational techniques are employed to predict the resulting optimal behaviors. This manner of examination may be called *normative* or *prescriptive*: a formulation of the task being solved is translated into prescriptions for how the nervous system *should behave* in order to solve the task. In contrast to this approach, many studies begin by modeling known properties of the nervous system. For instance, electrophysiology studies of the visual cortex may measure the firing properties of neurons in response to stimuli. These models are then used to describe how they give rise to properties such as Gabor-like receptive fields.[68] This approach may be called *descriptive*: measured properties are used to describe how the nervous system *does behave* when it solves tasks.

There are advantages and disadvantages to both approaches. Prescriptive models explain *why* the nervous system behaves as it does, but not *how* the behaviors are produced. Because prescriptive models are constructed to optimally solve the problems, they may not yield insight into *how* the nervous system's mechanisms produce behavior. However, descriptive models explain *how* mechanisms are responsible for the nervous system's behavior, but not *why* these mechanisms behave as they do. Moreover, relying on accurate knowledge of many measured parameters may render the predictions of descriptive models vulnerable to imprecision; in large, complex systems small deviations in parameters can result in gross and qualitative differences. Not surprisingly, the two methods are complementary and can inform each other. A prescriptive model of limb movements might, for instance, propose unreasonably large reflex gains certain to induce instability in a real limb. A descriptive study of spinal reflex pathways, however, would reveal the time delays a realistic model is saddled with. Thus, the two approaches help us to ask better, more informed questions in future motor control experiments.

CONCLUSION

Recognizing the inherent uncertainty in sensory information, many sensorimotor problems have been successfully described by applying the rules of Bayesian statistics to analyze their solutions. This same method holds promise for many of the outstanding problems in cue combination, motor control, cognitive science, and neuroscience in general. Bayesian statistics and normative modeling are complementary to descriptive studies of the nervous system. As such, combining the two approaches will provide models that not only predict, but also explain how the nervous system solves the problems it is confronted with. This insight into the purpose, and detail of structure, will provide a deeper understanding of the nervous system.

REFERENCES

1. Faisal AA, Selen LP, Wolpert DM. Noise in the nervous system. *Nat Rev Neurosci* 2008, 9:292–303.
2. Kording KP, Wolpert DM. Bayesian integration in sensorimotor learning. *Nature* 2004, 427:244–247.
3. Weiss Y, Simoncelli EP, Adelson EH. Motion illusions as optimal percepts. *Nat Neurosci* 2002, 5:598–604.
4. Miyazaki M, Nozaki D, Nakajima Y. Testing Bayesian models of human coincidence timing. *J Neurophysiol* 2005, 94:395–399.
5. Miyazaki M, Yamamoto S, Uchida S, Kitazawa S. Bayesian calibration of simultaneity in tactile temporal order judgment. *Nat Neurosci* 2006, 9:875–877.
6. Körding KP, Ku SP, Wolpert D. Bayesian Integration in force estimation. *J Neurophysiol* 2004, 92:3161–3165.
7. Knill D, Richards W. *Perception as Bayesian Inference*. Cambridge: University Press; 1996.
8. Chater N, Tenenbaum JB, Yuille A. Probabilistic models of cognition: where next? *Trends Cogn Sci* 2006, 10:292–293.
9. Peterka RJ, Loughlin PJ. Dynamic regulation of sensorimotor integration in human postural control. *J Neurophysiol* 2004, 91:410–423.
10. Tassinari H, Hudson TE, Landy MS. Combining priors and noisy visual cues in a rapid pointing task. *J Neurosci* 2006, 26:10154–10163.

11. Simoncelli EP, Adelson EH, Heeger DJ. Probability distributions of optical flow. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; Mauii, Hawaii: 1991, 310–315.
12. Ernst MO, Banks MS. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 2002, 415:429–433.
13. Alais D, Burr D. The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol* 2004, 14:257–262.
14. Maloney LT, Landy MS. A statistical framework for robust fusion of depth information. In: *Proceedings of the SPIE, Visual Communications and Image Processing IV*. Bellingham, WA: Pearlman; 1989, 1154–1163.
15. van Ee R, Adams WJ, Mamassian P. Bayesian modeling of cue interaction: bistability in stereoscopic slant perception. *J Opt Soc Am A* 2003, 20:1398–1406.
16. Kersten D, Yuille A. Bayesian models of object perception. *Curr Opin Neurobiol* 2003, 13:150–158.
17. Ernst MO, Bulthoff HH. Merging the senses into a robust percept. *Trends Cogn Sci* 2004, 8:162–169.
18. Geisler WS, Kersten D. Illusions, perception and Bayes. *Nat Neurosci* 2002, 5:508–510.
19. Rosas P, Wagemans J, Ernst MO, Wichmann FA. Texture and haptic cues in slant discrimination: reliability-based cue weighting without statistically optimal cue combination. *J Opt Soc Am A* 2005, 22:801–809.
20. Knill DC. Robust cue integration: a Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. *J Vis* 2007, 7:1–24.
21. van Beers RJ, Sittig AC, Gon JJ. Integration of proprioceptive and visual position-information: an experimentally supported model. *J Neurophysiol* 1999, 81:1355–1364.
22. Knill DC, Saunders JA. Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Res* 2003, 43:2539–2558.
23. Young MJ, Landy MS, Maloney LT. A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Res* 1993, 33:2685–2696.
24. Hillis JM, Watt SJ, Landy MS, Banks MS. Slant from texture and disparity cues: optimal cue combination. *J Vis* 2004, 4:967–992.
25. Hillis JM, Ernst MO, Banks MS, Landy MS. Combining sensory information: mandatory fusion within, but not between, senses. *Science* 2002, 298:1627–1630.
26. Jacobs RA. Optimal integration of texture and motion cues to depth. *Vision Res* 1999, 39:3621–3629.
27. Landy MS, Maloney LT, Johnston EB, Young M. Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Res* 1995, 35:389–412.
28. Flanagan JR, Bittner JP, Johansson RS. Experience can change distinct size-weight priors engaged in lifting objects and judging their weights. *Curr Biol* 2008, 18:1742–1747.
29. Berniker M, Kording K. Estimating the sources of motor errors for adaptation and generalization. *Nat Neurosci* 2008, 11:1454–1461.
30. Todorov E. Optimal control theory. In: Doya K, ed. *Bayesian Brain*. Cambridge, MA: MIT Press; 2006, 269–298.
31. Stengel RF. *Optimal Control and Estimation*. New York: Dover Publications; 1994.
32. Burge J, Ernst MO, Banks MS. The statistical determinants of adaptation rate in human reaching. *J Vis* 2008, 8:1–19.
33. Korenberg AT, Ghahramani Z. A Bayesian view of motor adaptation. *Curr Psychol Cogn* 2002, 21: 537–564.
34. Wolpert DM, Ghahramani Z, Jordan MI. An internal model for sensorimotor integration. *Science* 1995, 269:1880–1882.
35. Mehta B, Schaal S. Forward models in visuomotor control. *J Neurophysiol* 2002, 88:942–953.
36. Wallace MT, Roberson GE, Hairston WD, Stein BE, Vaughan JW, Schirillo JA. Unifying multisensory signals across time and space. *Exp Brain Res* 2004, 158:252–258.
37. Warren DH, Welch RB, McCarthy TJ. The role of visual-auditory "compellingness" in the ventriloquism effect: implications for transitivity among the spatial senses. *Percept Psychophys* 1981, 30:557–564.
38. Thurlow WR, Jack CE. Certain determinants of the "ventriloquism effect". *Percept Mot Skills* 1973, 36: 1171–1184.
39. Jack CE, Thurlow WR. Effects of degree of visual association and angle of displacement on the "ventriloquism" effect. *Percept Mot Skills* 1973, 37:967–979.
40. Kording KP, Beierholm U, Ma WJ, Quartz S, Tenenbaum JB, Shams L. Causal inference in multisensory perception. *PLoS ONE* 2007, 2:e943.
41. Sato Y, Toyoizumi T, Aihara K. Bayesian inference explains perception of unity and ventriloquism after effect: identification of common sources of audiovisual stimuli. *Neural Comput* 2007, 19:3335–3355.
42. Wei K, Kording KP. Relevance of error: what drives motor adaptation? *J Neurophysiol* 2008, 101: 655–664.
43. Wozny DR, Beierholm UR, Shams L. Human trimodal perception follows optimal statistical inference. *J Vis* 2008, 8:1–11.
44. Blakemore SJ, Frith CD, Wolpert DM. Spatio-temporal prediction modulates the perception of self-produced stimuli. *J Cogn Neurosci* 1999, 11:551–559.
45. Tenenbaum JB, Niyogi S. Learning causal laws. In: *25th Annual Conference of the Cognitive Science Society*; Cambridge, Massachusetts: 2003.
46. Steyvers I, Tenenbaum JB, Wagenmakers EJ, Blum B. Inferring causal networks from observations and interventions. *Cogn Sci* 2003, 27:453–489.

47. Gopnik A, Glymour C, Sobel DM, Schulz LE, Kushnir T, Danks D. A theory of causal learning in children: causal maps and Bayes nets. *Psychol Rev* 2004, 111:3–32.
48. Saxe R, Tenenbaum JB, Carey S. Secret agents: inferences about hidden causes by 10- and 12-month-old infants. *Psychol Sci* 2005, 16:995–1001.
49. Griffiths TL, Tenenbaum JB. From mere coincidences to meaningful discoveries. *Cognition* 2006, 103:180–226.
50. Holland P. Statistics and causal inference. *J Am Stat Assoc* 1986, 81:945–960.
51. Pearl J. *Causality*. New York: Oxford University Press; 2000.
52. Trommershäuser J, Maloney LT, Landy MS. Statistical decision theory and the selection of rapid, goal-directed movements. *J Opt Soc Am A* 2003, 20:1419–1433.
53. Maloney LT, Trommershäuser J, Landy MS. Questions without words: a comparison between decision making under risk and movement planning under risk. In: Gray W, ed. *Integrated Models of Cognitive Systems*. New York: Oxford University Press; 2006, 297–313.
54. Trommershäuser J, Gepshtein S, Maloney LT, Landy MS, Banks MS. Optimal compensation for changes in task-relevant movement variability. *J Neurosci* 2005, 25:7169–7178.
55. Najemnik J, Geisler WS. Optimal eye movement strategies in visual search. *Nature* 2005, 434:387–391.
56. Najemnik J, Geisler WS. Eye movement statistics in humans are consistent with an optimal search strategy. *J Vis* 2008, 8:1–14.
57. Najemnik J, Geisler WS. Simple summation rule for optimal fixation selection in visual search. *Vision Res* 2009, 49:1286–1294.
58. Kording KP, Wolpert DM. The loss function of sensorimotor learning. *Proc Natl Acad Sci U S A* 2004, 101: 9839–9842.
59. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. *Econometrica* 1979, 47: 263–291.
60. Trommershäuser J, Maloney LT, Landy MS. Decision making, movement planning and statistical decision theory. *Trends Cogn Sci* 2008, 12:291–297.
61. MacKay DJC. *Information Theory, Inference, and Learning Algorithms*. Cambridge, UK: Cambridge University Press; 2003.
62. Bishop C. *Pattern Recognition and Machine Learning*. Cambridge, UK: Springer; 2006.
63. Ma WJ, Beck JM, Latham PE, Pouget A. Bayesian inference with probabilistic population codes. *Nat Neurosci* 2006, 9:1432–1438.
64. Fiser J, Berkes P, Orban G, Lengyel M. Statistically optimal perception and learning: from behavior to neural representations. *Trends Cogn Sci* 2010, 14:119–130.
65. Angelaki DE, Shaikh AG, Green AM, Dickman JD. Neurons compute internal models of the physical laws of motion. *Nature* 2004, 430:560–564.
66. Harris CM, Wolpert DM. Signal-dependent noise determines motor planning. *Nature* 1998, 394:780–784.
67. Todorov E, Jordan MI. Optimal feedback control as a theory of motor coordination. *Nat Neurosci* 2002, 5:1226–1235.
68. Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 1962, 160:106–154.

