

# The Perception of Probability

C. R. Gallistel  
Rutgers University

Monika Krishan  
Rutgers University and India Ascension, Bangalore

Ye Liu  
Rutgers University and Chinese Academy of Sciences

Reilly Miller  
University College London and Princeton University

Peter E. Latham  
University College London

We present a computational model to explain the results from experiments in which subjects estimate the hidden probability parameter of a stepwise nonstationary Bernoulli process outcome by outcome. The model captures the following results qualitatively and quantitatively, with only 2 free parameters: (a) Subjects do not update their estimate after each outcome; they step from one estimate to another at irregular intervals. (b) The joint distribution of step widths and heights cannot be explained on the assumption that a threshold amount of change must be exceeded in order for them to indicate a change in their perception. (c) The mapping of observed probability to the median perceived probability is the identity function over the full range of probabilities. (d) Precision (how close estimates are to the best possible estimate) is good and constant over the full range. (e) Subjects quickly detect substantial changes in the hidden probability parameter. (f) The perceived probability sometimes changes dramatically from one observation to the next. (g) Subjects sometimes have second thoughts about a previous change perception, after observing further outcomes. (h) The frequency with which they perceive changes moves in the direction of the true frequency over sessions. (Explaining this finding requires 2 additional parametric assumptions.) The model treats the perception of the current probability as a by-product of the construction of a compact encoding of the experienced sequence in terms of its change points. It illustrates the why and the how of intermittent Bayesian belief updating and retrospective revision in simple perception. It suggests a reinterpretation of findings in the recent literature on the neurobiology of decision making.

*Keywords:* change-point representation, Bayesian model comparison, simplicity–adequacy trade-off, retrospective revision, compact coding

*Supplemental materials:* <http://dx.doi.org/10.1037/a0035232.supp>

We live in a world of nonstationary stochastic processes. Discrete events happen at some rate or with some probability. Most of these stochastic processes are nonstationary; their parameters are themselves stochastic variables; they change by unpredictable

amounts at unpredictable times. Under at least some circumstances, both human and nonhuman animals adapt their behavior to doubly stochastic processes in approximately optimal ways and with surprising quickness (Balci, Freestone, & Gallistel, 2009; Balci et al., 2011; Kheifets & Gallistel, 2012; Krugel, Biele, Mohr, Li, & Heekeren, 2009; Nassar et al., 2012; Nassar, Wilson, Heasley, & Gold, 2010). These results imply that the brains of both human and nonhuman animals have computational mechanisms that yield reasonably accurate perceptions of simple probabilities, even when those probabilities are not stationary.

The estimation of stepwise nonstationary stochastic parameters in real time is a nontrivial computational problem. Its solution has attracted recent interest from cognitive scientists and neuroscientists (Adams & MacKay, 2006; Wilson, Nassar, & Gold, 2010), as well as statisticians (Fearnhead & Liu, 2007; Polunchenko & Tartakovsky, 2012; Turner, Saatici, & Rasmussen, 2009). The computational complexity of the optimal (full Bayesian) solution is linear in the run length (Adams & MacKay, 2006) and exponential in the depth of the recursion (Wilson et al., 2010). This growth in computational complexity has been viewed as implau-

---

C. R. Gallistel, Rutgers Center for Cognitive Science, Rutgers University; Monika Krishan, Department of Psychology, Rutgers University, and India Ascension, Bangalore, Karnataka, India; Ye Liu, Rutgers Center for Cognitive Science, Rutgers University, and State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing, China; Reilly Miller, Gatsby Computational Neuroscience Unit, University College London, London, United Kingdom, and Princeton University; Peter E. Latham, Gatsby Computational Neuroscience Unit, University College London, London, United Kingdom.

We gratefully acknowledge support from Rutgers University and National Basic Research Program of China (2011CB302201).

Correspondence concerning this article should be addressed to C. R. Gallistel, Rutgers Center for Cognitive Science, Psychology Building Addition, Busch Campus, 152 Frelinghuysen Road, Piscataway, NJ 08854-8020. E-mail: [galliste@rucss.rutgers.edu](mailto:galliste@rucss.rutgers.edu)

sible from a neurobiological standpoint, motivating a search for algorithms that approximate the optimal computation with greatly reduced computational load. Most of those algorithms—the algorithms suggested as models for what brains are doing—use delta-rule updating to generate the estimate (percept) of the current probability (Brown & Steyvers, 2009; Krugel et al., 2009; Nassar et al., 2010, 2012; Steyvers & Brown, 2006).

Delta-rule updating algorithms have a long history in psychology, computer science, and neuroscience. The best known example in psychology is the Rescorla–Wagner model of associative learning (Rescorla & Wagner, 1972). Delta-rule updating is also widely used in reinforcement-learning algorithms in computer science applications (Sutton & Barto, 1998).

A delta-rule updating algorithm computes for each new event the discrepancy between the observed event and a prediction. The prediction is variously interpreted as a parameter of a stochastic process or as an associative strength or as the value of an act. An event that is perceived to have a low probability is generally predicted not to happen on any particular occasion. When it does happen, the prediction error, the discrepancy between prediction and occurrence, is large; when it does not happen, the error is small. Whether the event occurs or not, the estimate of its probability (the percept or the associative strength or the value) is adjusted by moving the preoccasion estimate toward the observed outcome by an amount that is some fraction of the prediction error. The fraction by which the estimate moves is the learning rate. The higher the learning rate, the faster the process adapts to changes, but the noisier and more inaccurate its estimates become. Put another way, delta-rule updating processes generate estimates of the current state of stochastic parameters by imposing a geometrically decaying running average window on the outcomes observed on previous occasions. The higher the learning rate, the narrower this window.

Delta-rule algorithms are recursive; that is, the prediction made on the next occasion is based on the updated estimate of the stochastic parameter(s) made on the previous occasion. In addition to their a long multidisciplinary popularity, delta-rule theories have at least four further attractions for behavioral, cognitive, and neuroscientists: (a) They are path independent, which means that future adjustments depend only on the current estimate and the next outcome. (b) Their computational complexity does not increase with the extent of the relevant experience (the length of the observed sequence). (c) They can be made to track stochastic parameters fairly well. (d) They appear neurobiologically plausible in the light of the discovery that dopamine neurons signal something like prediction errors (Dayan & Niv, 2008; Kakade & Dayan, 2002; Montague, Dayan, & Sejnowski, 1996; Niv, 2009; Schultz, Dayan, & Montague, 1997).

In nonstationary environments, the choice of a learning rate becomes an interesting problem. The problem is acute when these models are applied to the simplest and most common stochastic process, the Bernoulli process, whose parameter is  $p$ , the probability that on a discrete occasion one of two binary outcomes will be observed. When outcomes are binary, the outcome on any one occasion is always at one extreme or the other (0 or 1), but the hidden (unobserved) underlying probability parameter,  $p$ , is generally intermediate between those extremes. Because the observable outcomes are always at the extremes, delta-rule updating tugs the estimate of the probability around rather violently. To get a

smooth and accurate estimate of  $p$  requires a slow learning rate, but with a slow learning rate, the estimate adapts slowly to large step changes in  $p$ .

Results from several recently published experiments show that the apparent learning rate (i.e., the rapidity with which subjects' adjust their estimates of stochastic parameters under nonstationary conditions) varies over time (Brown & Steyvers, 2009; Nassar et al., 2010, 2012). To explain this, modelers have added to their algorithms processes for estimating volatility (Brown & Steyvers, 2009) or probability distributions on run lengths (Nassar et al., 2010, 2012). These estimates are used to modulate the learning rate in real time. In other words, subjects are assumed to learn faster on some occasions than on other occasions.

The relation between objective relative frequency (probability) and subjective probability is central to the enormous multidisciplinary literature on judgments and decisions. There is therefore a large experimental literature that aims at determining this relation (for reviews of the older literature, see Peterson & Beach, 1967; Rapoport & Wallsten, 1972; Slovic, Fischhoff, & Lichtenstein, 1977; Wallsten & Budescu, 1983). There is, however, almost no literature that is directly relevant to evaluating the adequacy of delta-rule models in explaining the human perception of nonstationary Bernoulli  $p$  values.

To assess the applicability of these computational models, we need experiments in which subjects estimate a stepwise nonstationary *descriptive* probability, not an *inferred* probability. This distinction was made by Peterson and Beach (1967) in their influential early review of the literature on probability estimation. It corresponds to the distinction in Bayesian statistics between parameter estimation and model selection.

In a descriptive-probability experiment, subjects are asked to estimate the parameter of a distribution after viewing draws from it. In the experimental literature, the to-be-estimated parameter is almost always either the  $p$  (probability) parameter of the Bernoulli distribution or the  $\mu$  (mean) of a normal distribution. These hidden parameters are continuous variables analogous to, say, light intensity. The task is similar to making brightness or loudness estimates, except that the estimate must be based on a sequence of one or more observations that are only stochastically related to the hidden quantity being estimated. The observed stimulus differs stochastically from the true parameter. There is, therefore, an issue about how to define the observed stimulus, an issue we address. A likely explanation for the dearth of experiments in which subjects estimated a Bernoulli probability after every observation is the difficulty of characterizing the objective stimulus to which the subject is responding. We call this “stimulus” the *observed probability*, to distinguish it from the hidden true probability.

What Peterson and Beach (1967) called an *inferred-probability* experiment might now be called a *relative-likelihood* experiment. It requires the subject to explicitly or implicitly choose between a small number of well separated possible source distributions, given some observed outcomes. In the simplest and most common case, there are only two possible source distributions. Subjects are told the values of their parameters, see some data (outcomes), and must then judge which distribution the data came from. Or, they may be asked to state the “probability” that the data come from one source or the other. Probability is in scare quotes because it is not the usual simple descriptive probability; rather, it is what is variously called the reverse, backward, or inverse probability—or most

commonly in recent years, the relative likelihood. Relative likelihood is an odds ratio; it ranges from 0 to infinity, unlike probability, which ranges from 0 to 1. A reverse probability is not the probability parameter of a Bernoulli distribution. If it were, then, as more and more outcomes are observed, estimates of it should converge on a value between 0 and 1, and this value should be a monotonically increasing function of the true  $p$ . By contrast, as more and more outcomes are observed in a reverse probability experiment, the relative likelihood should converge on 0 or infinity, depending on which choice (which source distribution) the odds favor. Put another way, when the true parameters of two source distributions,  $X$  and  $Y$ , have different values, then as the number of observed outcomes goes to infinity, the reverse probability that the source distribution is  $X$  goes to 0 or to 1, depending on whether  $X$  is in fact the source. Provided that the  $p$  parameters of the two distributions differ, then the reverse probability converges on one or the other limit, because with enough data there is always decisive evidence one way or the other.

Peterson and Beach observed that in descriptive probability estimation experiments, accuracy is generally the measure of how well a subject does, whereas in an inferred probability experiment optimality is generally the measure. Accuracy has commonly been measured simply by the degree to which subjects' estimates match the true value of the (hidden)  $p$  parameter. That approach, however, implicitly assumed that subjects could know that true value from their observation of a limited number of outcomes, which, obviously, they cannot. What we need to measure accuracy correctly is the ideal observer's descriptive probability, the best estimate that can be made given the observed outcomes. Optimality, by contrast, is measured by the degree to which the subjects' estimates conform to those of a statistical model of the inference and decision process. Peterson and Beach called the model of the ideal inferencer "statistical man." Statistical man tells us the ideal estimate of the reverse probability.

The distinction between a descriptive probability task and an inferred probability (reverse probability) task is critical, because, as Peterson and Beach (1967) pointed out, the two kinds of tasks consistently give quite different results as regards the relation between the objective probability and subjects' estimates of it:

The most striking aspect of the results [from descriptive probability experiments] is that the relation between mean estimates and sample proportions is described well by an identity function. The deviations from this function are small; the maximum deviation of the mean estimate from the sample proportion is usually only .03–.05, and the average deviations are very close to zero. (Peterson & Beach, 1967, p. 30)

Whereas they summarized the results from inference experiments as follows (their summary further explicates the concept of an inference experiment):

Imagine yourself in the following experiment. Two urns are filled with a large number of poker chips. The first urn contains 70% red chips and 30% blue. The second contains 70% blue chips and 30% red. The experimenter flips a fair coin to select one of the two urns, so the prior probability for each urn is .50. He then draws a succession of chips from the selected urn. Suppose that the sample contains eight red and four blue chips. What is your revised probability that the selected urn is the predominantly red one? If your answer is greater than .50, you favor the same urn that is favored by most subjects and

by statistical man. If your probability for the red urn is about .75, your revision agrees with that given by most subjects. However, that revised estimate is very conservative when compared to the statistical man's revised probability of .97. That is, when statistical man and subjects start with the same prior probabilities for two population proportions, subjects revise their probabilities in the same direction but not as much as statistical man does (Edwards, Lindman, & Phillips, 1965). (Peterson & Beach, 1967, p. 32)

In short, as early as 1967, it was clear that the function relating a simple probability to subjects' estimate of that probability is approximately the identity function, whereas the function relating an estimated reverse probability (inferred probability) to the true reverse probability commonly departs substantially from the identity. In this work, we are concerned with the perceptual process by which descriptive probabilities are formed, not the process by which reverse probabilities are formed.

Another important methodological aspect of the kind of experiment required to test delta-rule updating models is that subjects report their estimate of the parameter rather than make a prediction. Experiments asking for predictions are much more common than experiments asking for parameter estimates. Experimenters often attempt to infer the subjects' parameter estimates from their predictions. But doing so requires a theory about the relation between those estimates and the predictions based on them. When one asks for the estimates directly, no inference back to an estimate is required, though there is, of course, an issue about the relation between subjects' behaviorally expressed estimate and their underlying "true" estimate. This is a vexing issue in much of the psychophysical literature (for example, the magnitude estimation literature). However, it proves not to be a problem for direct estimates of probability because, as noted above by Peterson and Beach (1967), the function relating the behaviorally expressed descriptive probability to objective probability usually approximates the identity.

Recent work by Brown and Steyvers (2009) illustrates the distinction between descriptive and inferred probability and the methodological difference between asking for a description of the perceived probability versus for a prediction based on it. Their subjects viewed stepwise nonstationary sequences in which the source of the sequentially observed data changed on occasion. Subjects were asked on every trial both to predict the next observation and to say which source they believed the most recent observation came from. Subjects were told the two distributions' parameter values at the outset, and these values did not change during the experiment. What changed stochastically was the source distribution. Subjects had to decide which was the more likely source distribution for the currently observed data and make a prediction based on that decision. Like Brown and Steyvers, we would assume that their decision was based on a computation of an approximation to the Bayesian relative likelihood, aka the reverse probability.

In summary, the results we seek to model must come from descriptive probability experiments in which the true probability is stepwise nonstationary and subjects make known their estimate observation by observation. Only under these conditions can we observe the transient responses of the perceptual system to abrupt changes in the input. These transient responses are critical for what an engineer would call system identification.

Despite helpful suggestions from many sources, we have found only one published experiment of the kind needed to assess the applicability of delta-rule models to the perception of stepwise nonstationary simple Bernoulli probability (Robinson, 1964). The results pose deep difficulties for any kind of trial-by-trial updating model, including a fortiori, delta-rule models. Trial-by-trial updating models cannot explain the statistics of the step-like pattern of estimates seen in all subjects: Subjects' behaviorally expressed trial-by-trial estimates often stay constant for many trials (wide step widths). Of importance, the estimates sometimes jump by very large amounts from one trial to the next (large step heights). However, equally important is the fact that steps of all sizes are common and that the smallest steps are the most commonly observed. We argue that the joint distribution of step widths and step heights is not consistent with a trial-by-trial updating model.

We have replicated and extended Robinson's results. These results motivate our new and quite different computational model for the process that mediates the perception of a descriptive probability. This model also suggests a reinterpretation of the results from recent neuroimaging experiments.

Robinson reported results from 4 "regular" subjects and 4 pilot subjects. We report comparable results from 10 more subjects. Robinson's experiment and ours were done in the psychophysical tradition, in which each subject is a statistical universe. In a psychophysical experiment, one runs enough trials on each subject—10,000 in our case—to be able to say with confidence what the truth is in that subject. One runs additional subjects not to obtain enough data to make good estimates based on cross-subject averages but rather to determine whether the results obtained from one subject may be replicated with another. The experimental results that challenge delta-rule models and motivate our model were obtained from each of Robinson's 8 subjects and from each of our 10 subjects. In other words, these results have been replicated 18 times in 18 different subjects, in two different labs, almost 50 years apart, with methods that differed in many of their details. These results are sufficiently secure to place strong constraints on computational models of the process by which the typical human brain computes a hidden Bernoulli probability.

### Robinson's Experiment

Robinson's subjects used a lever controlling a dial to indicate a continuous estimate of the hidden parameter,  $p$ , of the stepwise nonstationary Bernoulli process that controlled the flashing of the two lights above the indicator dial (see Figure 1). The probability that a flash would be on the right was  $p$ ; the probability that it would be on the left was  $(1 - p)$ . The lever and associated mechanisms contained only enough coulomb friction to retain a setting without constant force. Neither springs nor viscous frictions were used.

Five flash rates were used, ranging from .5 to 8 flashes per second in octave steps. The probability,  $p$ , changed in discrete steps at irregular intervals. There were 8 different heights for these steps. They were arranged in two separate problem sets (blocks of steps), a large-step problem set and a small-step problem set. In the large-step problems, the absolute values for step size  $|p_a - p_b|$ , where  $p_a$  is the probability after the step and  $p_b$  the probability before the step, were .16, .32, .48, or .64. In the small-step problems, they were .06, .12, .18, or .24. Step widths—the number

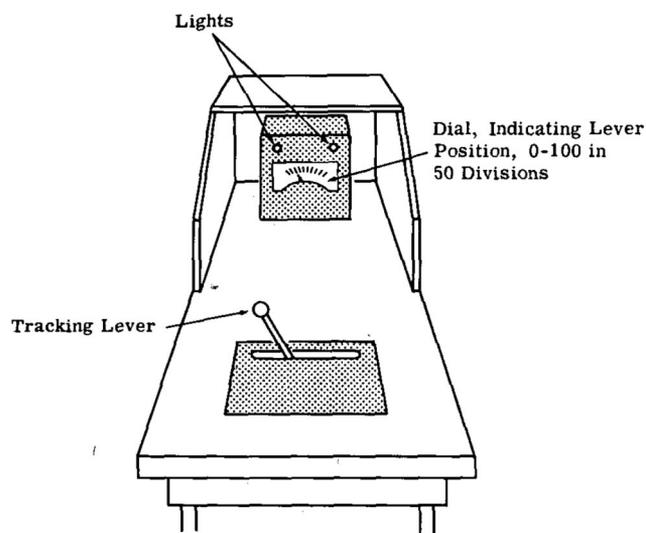


Figure 1. Robinson's apparatus for real-time tracking of a stepwise nonstationary Bernoulli process governing the flashing of the two lights above the dial that indicated the lever position. The subjects kept their hand on the lever at all times, adjusting its position (hence also the needle on the dial) to indicate their current perception of the hidden parameter that governed the relative frequency with which the two lights flashed. Lever position was recorded digitally (on punch cards; this was the early 1960s) and on a chart recorder. From "Continuous Estimation of Time-Varying Probability," by G. H. Robinson, 1964, *Ergonomics*, 7, p. 9. Copyright 1964 by Taylor & Francis. Reprinted with permission.

of flashes between changes in  $p$ —ranged from 34 to 89 flashes. Seventeen values of  $p$  were used. The large-step problems used values of .02, .18, .34, .5, .66, .82, and .98. The small-step problems used values of .08, .14, .26, .32, .44, .5, .56, .68, .74, .86, and .92.

Flash rates and small- and large-step problem sets were exhaustively combined in a random order. In each experimental session, a subject saw two or three problem sets, presented in blocks, each with a particular rate and step range. There were 4 pilot subjects and 4 "full-scale" subjects. Each of the latter performed the task in 15 sessions of about an hour each. They were paid at a fixed rate minus their accumulated squared error. In other words, the more accurately they performed, the more they earned. They were instructed in detail on the concept of time-varying probability and the estimation requirements of the task but were not told that the changes would be step changes nor how big or how frequent these would be. They were told that the frequency of the changes would vary. They got 45 minutes of practice before the regular sessions began. Pilot work showed that the response form changed little after 30 minutes practice and that subjects failed to recognize or improve their performance with the repetition of identical problems.

Robinson summarized his results as follows:

The general response form shown in Figure 2 was found in 80 to 90 per cent of the 3440 step responses that were analyzed. The response was characterized by rapid changes separated by period of little or no change in the estimate. This response form suggests that the task behavior might be described in terms of a series of decisions concerning possible changes in the probability. (Robinson, 1964, p. 11)

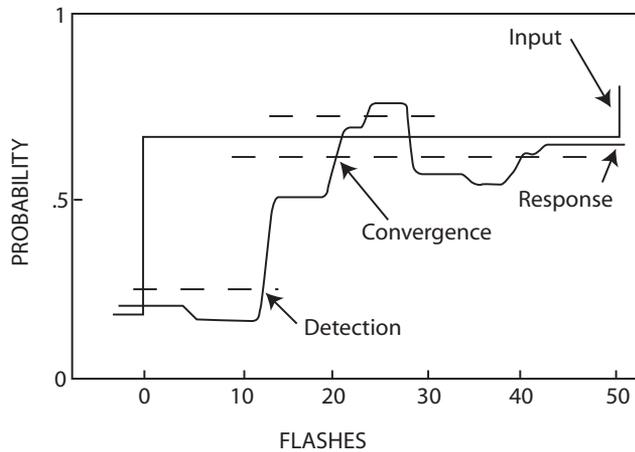


Figure 2. Lever setting as a function of time and flash number during a typical change in probability,  $p$ . Note the steplike (discontinuous) nature of the adjustment, both when a change is detected and when subsequent adjustments are made to bring the new estimate into register with the new value of  $p$ . The model we propose reproduces this behavior, whereas delta-rule updating models cannot be made to reproduce it (see Figure 12). Convergence is the point at which the lever setting crossed the dashed limits of “acceptable accuracy” around the true value. Detection is the first move of the lever in the correct direction following a change in the true  $p$ . Reprinted from “Continuous Estimation of Time-Varying Probability,” by G. H. Robinson, 1964, *Ergonomics*, 7, p. 11. Copyright 1964 by Taylor & Francis. Reprinted with permission.

He also noted that “no qualitative differences appeared between the eight subjects observed in the main and pilot studies” (p. 12).

Robinson measured the Detection latencies, the Convergence latencies (see Figure 2), the accuracy (mean error), and the precision (root mean square error) after convergence. The accuracy measure showed the mapping from true probability to average reported probability to be the identity function: “No significant mean error was measured at any probability level with any combination of flash rate, step size . . . or subject. The largest error observed was smaller than the least scale division on the subject’s response indicator (0.02)” (Robinson, 1964, p. 15). The precision, as measured by the root mean square error, was approximately what one would expect from a running average of the last 17 flashes.

Detection latencies were short. How short depended, as one would expect, both on the size of the step—big steps were detected faster than small ones—and on the flash rate: at higher rates more flashes occurred between a step change and its detection (i.e., the answering change in lever position). At the slower rates, one flash every 1 or 2 seconds (which is the self-paced presentation rate chosen by the subjects in our replication of this experiment), a step of .1 was detected after 10–15 flashes, whereas a step of .6 was detected after 4 to 5 flashes.

Of importance, detection latencies also depended on the problem set during which a step of a given size occurred:

In the region of overlapping step size (0.16 to 0.24) the small step problem yields detections of from one to six flashes higher than the large step problem at all rates. The subjects were evidently modifying their tracking method according to the type of problem being pre-

sented. The subjects had no prior indication, of course, that there were two problem types. The subjects appear to have made larger, more decisive, response changes on the large step than on the small step problem. This change to a more responsive behavior is appropriate in quickly reducing the large errors following the larger step changes. (Robinson, 1964, p. 13)

In other words, Robinson’s subjects appeared to form problem-set-specific priors on step size, even though they initially did not know that the steps were grouped on the basis of size into two distinct problem sets (a small-step set and a large-step set) and even though the sets were randomly intermixed, with no signal at the boundary between sets. This finding implies that people do not simply track current probability; rather, they extract from experienced sequences a representation of the size of the steps they have experienced.

That subjects encode and remember the steps they experience is the central idea in the model of probability perception that we propose. In our model, the current perception of the hidden  $p$  is a by-product of a process that constructs a compact encoding of experienced sequences, encoding them by their change points. The size and frequency of the experienced changes are trivially extractable from such a record. Extracting this information provides an informative prior for the processing of current input.

Robinson prefaced his presentation of a descriptive model as follows:

The general response form suggests that the subject was making a series of decisions concerning possible changes in the probability. These decisions in turn led to fairly sudden response adjustments. This mode of behavior is undoubtedly partially derived from the step function nature of the input. It was also observed, however, in an informal experiment using a slow, continuously changing probability.

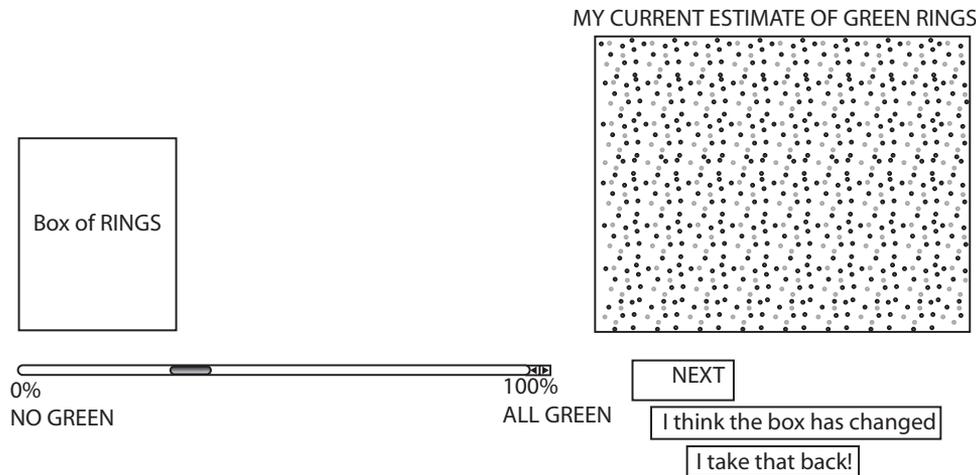
A descriptive model for this task must account for three phenomena: the static estimation is performed about as well as a simple 17 sample averaging model; there is a minimum change, or perhaps rate of change, in the probability that can be detected; and when the subject decides that a change has occurred his response is essentially discontinuous. (Robinson, 1964, p. 16)

We drew the same conclusions from our replication of his experiment. At the time we ran the experiment and drew these conclusions, we were unaware of Robinson’s experiment. These conclusions motivated our own model.

## Our Replication

In our realization of the experiment, subjects view on a computer monitor the display shown in cartoon form in Figure 3, and they use a mouse to click on the *Next* button to draw rings from the “box of RINGS.” They are asked to use the mouse to adjust the slider beneath the box to indicate their draw-by-draw estimate of  $p_g$ , which denotes the proportion of green circles in the box. Thus, they are not estimating the observed proportion but rather the unobserved proportion of circles in the box. Of course, their estimate of the hidden proportion must be based on the observed proportion, but it may depart from it quite markedly when the number of observed circles is very small; as, for example, at the beginning of a session.

Moving the slider varies the proportion of green and red circles in the large box at the right of the screen, which contains 1,000



*Figure 3.* Cartoon of the screen seen by the subjects. In the feedback box at the right, red circles are black and green circles are gray. Clicking on the *NEXT* button caused another circle to emerge from the box of RINGS and then evaporate. Subjects reported a perceived change in  $p_g$  (the probability of a green circle emerging) by clicking on the “I think the box has changed” button. They reported second thoughts about their most recent change report by clicking on the “I take that back!” button. The subject’s perception of  $p_g$  on each trial was taken to be the slider position recorded on that trial.

circles. The function mapping from slider position to the proportion of green circles in this box is the identity. This visual feedback may initially be helpful, but most subjects report that they pay little or no attention to it once they get into the task. Subjects are told that the box of RINGS may sometimes be silently replaced by a box with a different proportion of green circles. They are asked to report a perceived change in  $p_g$  by clicking on the “I think the box has changed” button. When debriefed, the first 5 subjects spontaneously reported having second thoughts about some of their change reports: After seeing more circles, they sometimes felt that there had not, after all, been a change. Because we believe this reveals an important aspect of the process that mediates the perception of probability, we added the “I take that back!” button. We told the last 5 subjects to click on it to report a second thought.

Our subjects were allowed to practice the task until they felt comfortable with it, which took about 10 minutes. Each of our 10 subjects then completed 10 sessions, with 1,000 trials (draws from the box of RINGS) per session, working at their own pace. The instructions we gave them and their demographics are in [Appendix A](#). They were paid \$10 per session. The quality of their performance had no effect on their remuneration, whereas Robinson’s subjects were penalized on the basis of their root mean square error.

Sessions lasted approximately half an hour ( $M = 25.6$  min; standard deviation = 6.1), which means that subjects chose to draw from the box at a rate of about 1 circle every 1.2 to 1.8 seconds (the slow end of Robinson’s flash rates). The programmed values for the proportion of green circles in the box were roughly evenly distributed between 0 and 1 (see [Figure 4A](#)). The size and direction of a change in this hidden parameter were determined by a random choice of the next value from a uniform distribution between 0 and 1, subject to the restriction that the resulting change in the odds,  $p_g/(1-p_g)$ , be no less than fourfold. The sizes of the experienced changes in  $p_g$  were roughly evenly distributed within the intervals  $-.85$  to  $-.2$  and  $.2$  to  $.85$  (see [Figure 4B](#)).

Further details of our method are in [Appendix A](#). The MATLAB code for our implementation of the experiment is in the online supplemental material.

Robinson’s version and our version of the experiment were run in different laboratories almost 50 years apart. We were unaware of his experiment when we ran ours. The results of theoretical importance were obtained from all 8 of Robinson’s subjects and from all 10 of ours. Given the difference in apparatus, the difference in the method of indicating perceived probability, the fact that Robinson’s subjects’ remuneration depended on the accuracy of their performance, whereas ours did not, and the many differences in the details of the procedures, it appears that the results from this very simple experiment are robust to variation in particulars of the method. [Robinson \(1964\)](#) remarked in this regard, “The speed of response to the changes in probability was also faster than had been anticipated as was the consistency among the subjects. The consistency is particularly surprising considering the lack of instruction on the dynamic form and the ambiguity of a task requiring both detection and smoothing or averaging” (p. 19).

The probability that a change in  $p_g$  would occur after any given trial, denoted  $p_c$ , was constant at .005. Thus, the number of trials between changes was geometrically distributed, with an expectation of 200 trials between changes. Sometimes there were very few trials between changes (as few as 1) and sometimes as many as several hundred. In Robinson’s experiment, on the other hand, the number of flashes between changes varied over a rather narrow range, from 34 to 89 flashes.

## Results From Our Experiment

**Stepped updating.** Robinson’s subjects held a low-friction lever to move a dial, whereas ours used a mouse to move a slider on a computer screen. Nonetheless, we observed in all 10 of our subjects the same pattern that Robinson observed in all 8 of his (see [Figure 5](#)).

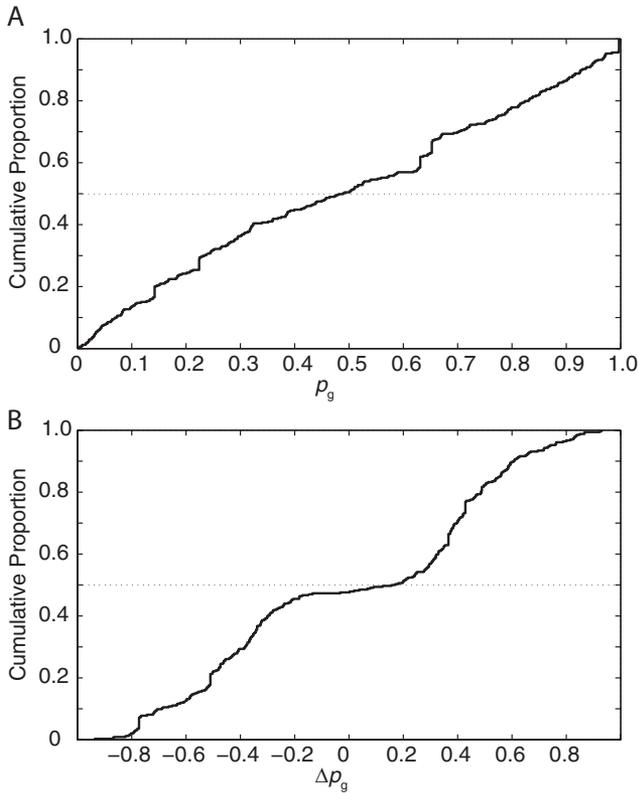


Figure 4. A: The cumulative distribution of  $p_g$  across subjects and sessions. B: The cumulative distribution of the signed magnitude of the step changes in  $p_g$  across subjects and sessions.

**Mapping is the identity function.** Like Robinson and in accord with the Peterson and Beach summary of the relevant literature up to 1967, we find that the mapping from true probability to median reported probability is the identity. The median trial-by-trial slider settings track closely the hidden true probabilities (see Figure 6). This is consistent with Robinson's finding that there was no significant mean error at any value of the probability.

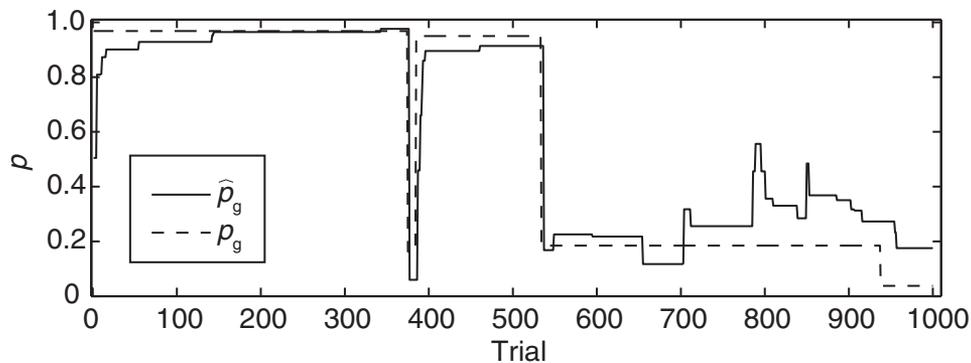


Figure 5. The trial-by-trial true probability of a green circle (denoted  $p_g$ , dashed line) and the trial-by-trial slider setting (denoted  $\hat{p}_g$ , solid line) for Subject 4, Session 8. Note the step-hold pattern of slider adjustment. This pattern is seen in every session from every subject. The resemblance to Robinson's example (Figure 2) is striking.

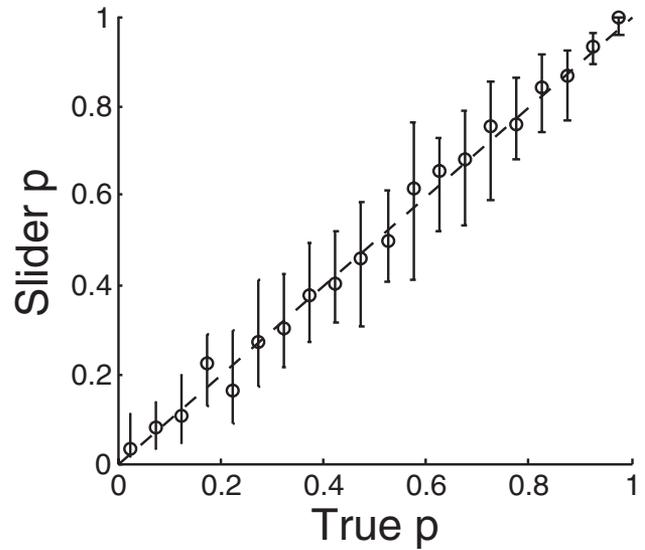


Figure 6. Median slider setting (open circles) and interquartile intervals (vertical lines) plotted against the bin centers (.025, .075, etc.), after binning of the data at intervals of 0.05.

**Precision.** The precision with which subjects estimate the probability is the measure of their average trial-by-trial error. However, the appropriate measure of error requires some discussion, because, as noted in our introduction, in measuring precision, one should use the "observed" probability, not the hidden, unobserved true probability. To anticipate, when appropriately measured, subjects' precision of estimation is the same at all values of the observed probability, except the most extreme (see Figure 18).

**Rapid detection of changes.** Like Robinson's subjects, our subjects detected changes quickly (see Figure 7). Our measure of the change-detection latency in slider setting is the number of trials between a true change and the first appropriately signed step in the slider setting thereafter. This is the same as Robinson's measure. Our measure of the expressed change latency is the number of trials between a true change and the first click on the "I think the

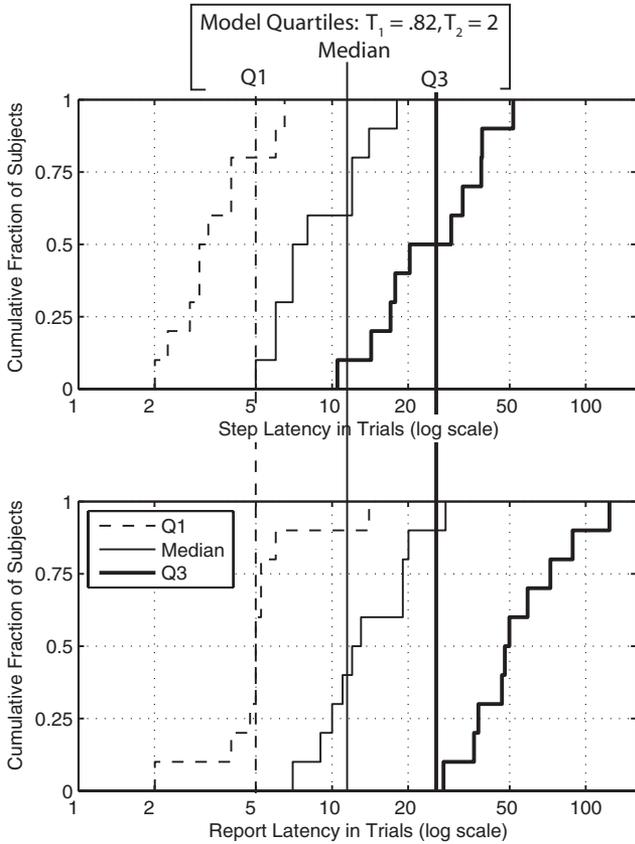


Figure 7. Cumulative distributions of the quartiles of the distributions of subjects' change-detection latencies. Top panel: The thin dashed line is the distribution across the 10 subjects of each subject's first quartile latency to make a slider adjustment in the appropriate direction in response to an objective change in  $p_g$ . Only adjustments made before the next change are tabulated. The thin solid line gives the distribution of the subjects' median latencies, and the heavy solid line gives it for the third quartile of their latencies. The median latency of the median subject was about 10 trials. Median slider adjustment latency is the same measure as Robinson's measure of Detection Latency, and it gives similar values. Largish changes are detected very quickly. Bottom panel: The distribution of latencies to report a change (latency to click on the "I think the box has changed" button). These were slightly longer than the latencies to respond with an appropriate slider adjustment. This may be in part or entirely an artifact of the fact that "false alarm" slider adjustments were more frequent than "false alarm" reports. The vertical lines labeled Q1, Median, and Q3 are the median quartiles predicted by the change-point model (described in detail below) when its decision criteria were set to  $T_1 = .82$  and  $T_2 = 2$ . That is, the thin vertical dashed line is the model's median first quartile across the 10 subjects, the heavier vertical black line is the model's median across the 10 subjects, and the heavy black vertical is the model's median third quartile across the 10 subjects.

box has changed" button. Figure 7 also contains predictions of our model; these are described in detail later.

**High hit rates and low false alarm rates.** Robinson did not compute hit and false alarm rates, but we did (see Figure 8). We take the hit rate to be the number of intervals between two true changes in which the subject clicked "I think the box has changed" at least once divided by the number of true changes. (Some true

changes were in principle undetectable because they lasted only one or two trials.) We take the false alarm rate to be the number of extra times the subject clicked "I think the box has changed" (i.e., the number of clicks after the first correct click and before the next change) divided by the number of trials on which a change call would have been scored as a false alarm. Thus, if subjects called a change after every trial, their false alarm rates would be 1; if they called a change at most once in any span between two true changes, their false alarm rate would be 0. Obviously, there are many more opportunities to make false alarms than to make hits. This is inherent in change detection. Nine of the ten subjects had mean hit rates in the range .77–1 and mean false alarm rates in the range .0004–.02 (dashed box in Figure 8).

**Second thoughts.** Subjects sometimes have second thoughts about their most recent perception of a change. After observing more outcomes, they decide that their most recent previous decision that there had been a further change is no longer justified (see Figure 9).

**Results Summary**

The results from Robinson's experiment and our replication of it place the following constraints on the computation that generates the percept of the probability,  $p$ , of a hidden and nonstationary Bernoulli process:

- It delivers reasonably accurate estimates of probability over the full range of its possible values, not systematically over or underestimating anywhere in this range. See Figure 6 above.
- It explicitly detects changes (as opposed to simply adapting to them). See Figure 7 above.

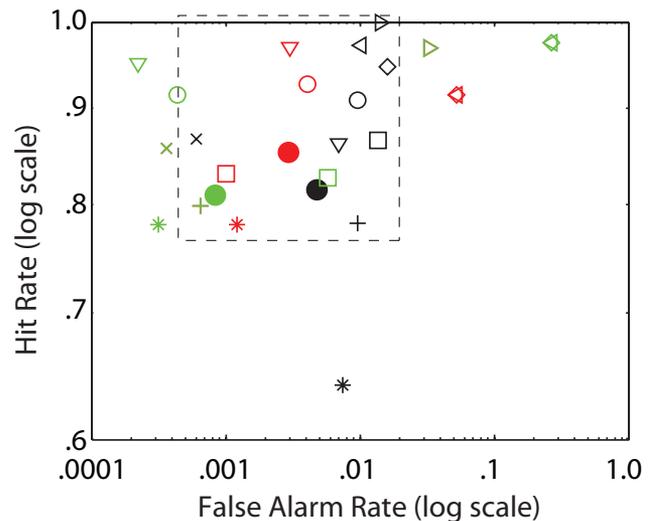


Figure 8. Mean hit rates versus mean false-alarm rates for subjects (black symbols) and for the model's simulation of each subject (the red and green symbols) in double-logarithmic coordinates. The symbols from the model's simulation have the same form as the corresponding subject symbol; the red symbol of a given form approximately minimizes the simulated false alarm rate for that subject, and the green one approximately minimizes the simulated hit rate. The light dashed square encompasses all but one subject. Note that the model symbols fall in and around this square. See later text for details on the estimation of the model parameters,  $T_1$  and  $T_2$ .

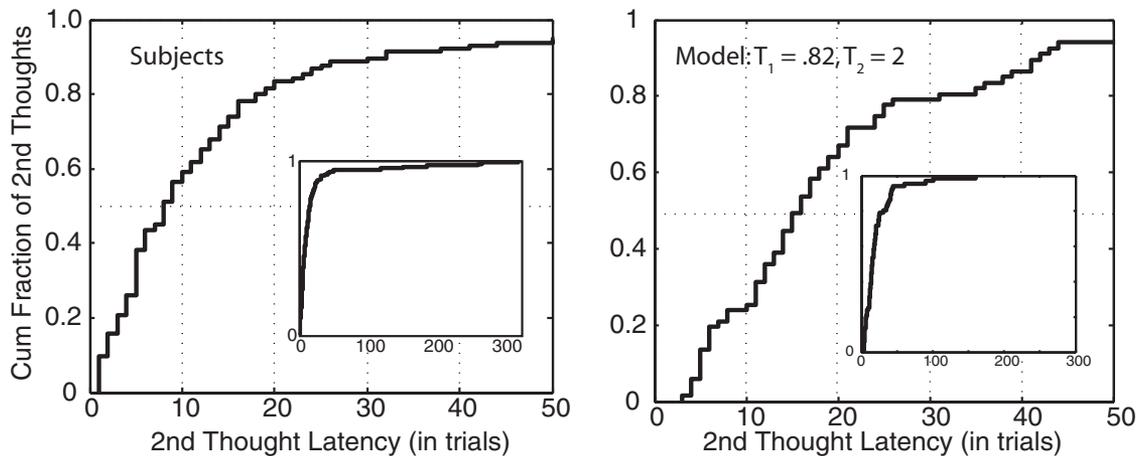


Figure 9. The cumulative (Cum) distribution of second thought latencies from subjects (left panel) and from the model (right panel; parameter estimation described later in text); 95% of the subjects' second thought latencies were fewer than 50 trials, but there was a very long tail; inset shows the complete distribution.

- It quickly detects substantial changes in probability, with a high hit rate and a low false-alarm rate. See Figure 8 above.
- It can make large discontinuous changes in the estimate between one observation and the next. See Figures 2 and 5 above and Figure 11 below.
- It can maintain a single, unvarying estimate for many trials. See Figure 5 above and Figure 11 below.
- It explicitly distinguishes between, on the one hand, errors in the estimate of  $p$  that become apparent as more data accumulate since the last change and, on the other hand, errors that arise because the underlying probability of the Bernoulli process has changed again. This is required by the fact that subjects often change the slider position without reporting a change.
- It modifies its estimate of how frequent true changes are (volatility), and this affects the rate at which it detects changes. See Figure 16 below.
- It has second thoughts regarding the validity of a previously perceived change; that is, it sometimes decides later that there was not in fact a change. See Figure 9 above.

### Theoretical Development

Our theoretical developments proceed as follows: We show first the pattern of trial-by-trial estimates produced by delta-rule models. It differs strikingly from the pattern subjects generate. We then consider whether a delta-rule model—or any trial-by-trial updating model—can be made to generate patterns like those generated by subjects if we assume that there is an output threshold. We conclude that they cannot be made to do so. There does not seem to be a way to get a trial-by-trial updating model to reproduce even approximately the joint distribution of step widths and step heights produced by subjects. We then develop our model, which is a Bayesian change-point estimation model with evidence triggered updating. We explain why the model is computationally efficient. We describe how we determine values for its parameters. We compare its performance to the performance of our subjects, and we conclude by discussing the implications of our model and the features of it that are of more general interest.

The two basic ideas behind the model are simple: First, the brain does not update its estimate of probability trial by trial; it updates only when there is evidence that its current estimate is unsatisfactory. Second, when it decides that the current estimate is unsatisfactory, it does Bayesian model comparison to decide between three different explanations: (a) the current estimate was based on a misleading set of observations, (b) there has been a change in the true value of  $p$ , or (c) the last change point added to the representation of the experienced sequence was unjustified in the light of later observations.

In this model, the perception of the current  $p$  is a by-product of the construction of a compact representation of the observed sequence in terms of its change points. When there is a problem with the current estimate, a new one is made. The observations on which the new estimate is based vary depending on which of these three explanations for the problem is most likely. Generating a compact model of the sequenced experienced so far provides a basis for anticipating and interpreting subsequent experience in the same context.

### The Problem for Trial-By-Trial Updating Models

If the step-hold pattern of slide adjustments accurately tracks the underlying percept, then that percept is not generated by the kind of trial-by-trial delta-rule updating process that has been widely assumed in the modeling literature. Figure 10 shows the behavior of the model proposed by Behrens, Woolrich, Walton, and Rushworth (2007) for the estimation of Bernoulli probability. (They modeled the process but did not obtain estimates of descriptive probabilities from their subjects.) This is representative of a popular class of models of the process by which the brain estimates nonstationary stochastic parameters (Cohen, McClure, & Yu, 2007; Corrado, Sugrue, Seung, & Newsome, 2005; Glimcher, 2003a, 2003b, 2009). In these models, the current estimate of the probability and, in some models, the current estimate of the volatility (the variance in the hidden probability parameter) are updated trial-by-trial. The volatility estimate modulates the probability-learning rate, which is the rate at which the estimate of

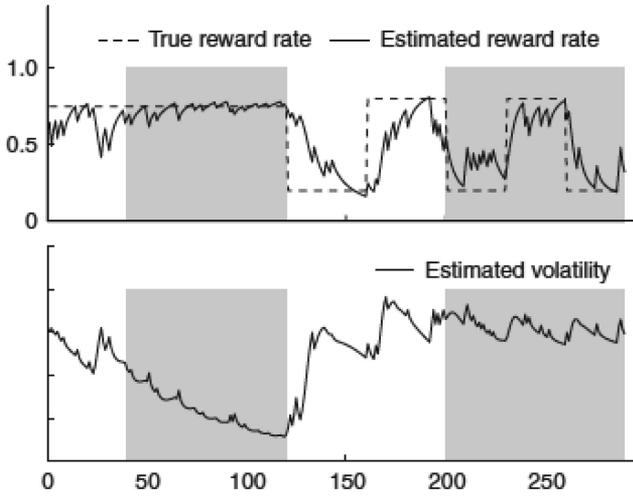


Figure 10. The behavior of the model proposed by Behrens et al. (2007) for the process by which the brain estimates rates of reward in their probability tracking task. The upper panel shows the model’s trial-by-trial estimate of the probability that choosing a blue square will be rewarded (in a binary choice task with differential reward). The estimate of this probability is updated on every trial, because it is a running average of the recent outcomes and those outcomes are binary. Because it is a running average, it adjusts relatively smoothly to changes in the probability that it is estimating. The rate at which it adjusts (the learning rate) is governed by the model’s estimate of the volatility (lower panel). The higher this estimate, the faster a new probability is learned. The estimate of volatility is also adjusted trial-by-trial. From “Learning the Value of Information in an Uncertain World,” by T. E. J. Behrens, M. W. Woolrich, M. E. Walton, and M. F. S. Rushworth, 2007, *Nature Neuroscience*, 10, p. 1215. Copyright 2007 by Nature Publishing Group. Reprinted with permission.

the probability adjusts to a change. Kalman filters are an example of and, often, an inspiration for such models. They are widely used in engineering to track nonstationary dynamic systems with noisy outputs or measurements thereof.

### Adding an Output Threshold Does Not Solve the Problem

Because trial-by-trial updating is a natural and pervasive first assumption, we investigated whether it is possible to explain the observed pattern of lever or slider settings on the assumption that the underlying percept really is updated trial by trial but subjects only reposition the lever or the slider when their percept differs by a greater-than-threshold amount from the current position. In other words, the step-hold pattern is seen because subjects are reluctant to make small changes in the observable output. This is the assumption that Robinson made in his descriptive model, but he did not test it against appropriate statistics computed from his records of lever position.

We considered a running average model, much like that proposed by Behrens et al. (2007) but with a difference threshold on the output. In this model, on each trial the current observable estimate (the slider setting),  $\hat{p}_g$ , is compared to an (unobserved) running average, denoted  $p_{ra}$ . If the absolute difference,  $|\hat{p}_g - p_{ra}|$ ,

exceeds a threshold, then  $\hat{p}_g$  is set to  $p_{ra}$ ; otherwise, it is not changed.

A fixed threshold suppresses all steps whose height is less than the threshold. This eliminates small step heights and makes the distribution of signed step heights bimodal. The data, however, show essentially opposite features: Small step heights are the most common, and the distribution of step heights is unimodal (see Figure 11). To try to remedy this, we used a variable (stochastic) threshold: on each trial, the threshold was drawn from a Gaussian distribution with mean  $\mu_\tau$  and coefficient of variation,  $cv_\tau$  ( $\leq .33$ ). If on any trial  $|\hat{p}_g - p_{ra}| > \mathcal{N}(\mu_\tau, cv_\tau \mu_\tau)$ ,  $\hat{p}_g$  was automatically updated. ( $\mathcal{N}(\mu_\tau, cv_\tau \mu_\tau)$  denotes a random draw from a normal distribution with mean  $\mu_\tau$  and standard deviation  $cv_\tau \mu_\tau$ .)

For the simplest running average model,  $p_{ra}$  is updated using a single exponentially decaying kernel,

$$p_{ra}(t) = (1 - \lambda)p_{ra}(t - 1) + \lambda x(t),$$

where  $t$  denotes trial,  $x(t) = 1$  if there is a green circle on trial  $t$  and 0 if there is a red one, and  $\lambda$  controls how far into the past the average extends, with smaller  $\lambda$ s corresponding to longer memories (and thus slower updating).

The main advantage of this model is its simplicity. However, it is really appropriate only when the true probability of a green circle,  $p_g$ , changes slowly, since in that case the decay rate,  $\lambda$ , can be matched to the rate at which  $p_g$  changes. It is not surprising, therefore, that this model was not a good match to the data. In particular, it could not produce the numerous large steps that one sees in the marginal distribution of subjects’ step heights in Figure 11.

To remedy this, and to explicitly take into account the sudden changes in  $p_g$ , we considered a more complex, two-kernel model. This model kept track of two running averages:  $p_{ra,fast}$ , which had

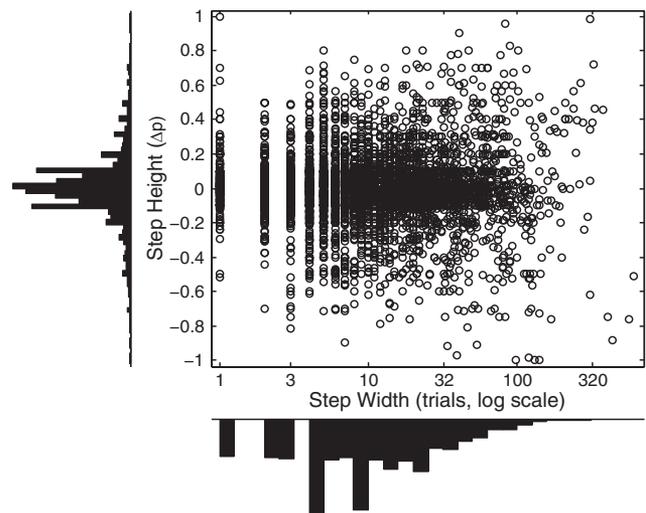


Figure 11. The joint distribution of subjects’ step widths (x-axis, log scale) and step heights (y-axis, linear scale) together with the marginal distributions. As may be seen from the marginal distribution of step heights, more than 50% of step heights fall in a narrow range around 0. In other words, the smallest possible movements of the slider are the most common, which they cannot be if there is a difference threshold on the output that suppresses small steps.

a short memory (large  $\lambda$ ), and  $p_{ra,slow}$ , which had a long memory (small  $\lambda$ ). The idea is that when there is a sudden change in  $p_g$ , the short memory running average will react much more quickly than the long memory average and that will signal a change. We thus introduced a second threshold, denoted  $\Delta_c$ . On each trial we computed  $|p_{ra,slow} - p_{ra,fast}|$ . If this was below  $\Delta_c$ ,  $p_{ra}$  was set to  $p_{ra,slow}$ ; if it was above  $\Delta_c$ ,  $p_{ra}$  was set to  $p_{ra,fast}$ . Updating of  $\hat{p}_g$  was the same as in the simple, one-kernel model: it was updated if  $|\hat{p}_g - p_{ra}|$  exceeded a (stochastic) threshold; otherwise it wasn't.

The two-kernel model adds two free parameters (a second decay rate and  $\Delta_c$ , the threshold that causes a switch from the slow to the fast running average), making a total of five free parameters. However, none of these additions to the complexity of the trial-by-trial updating model appreciably diminish the bimodality in the distribution of step heights. In fact, at every parameter combination we have tried, there were two well-separated modes in the distribution of step heights (see Figure 12).

We conclude that a model that updates trial by trial cannot approximate the distribution of step widths and heights that subjects produce. Thus, the joint distribution of step widths and step heights in Figure 11 is a strong constraint on viable models. Moreover, even if a running average model could reproduce the distribution of step widths and heights, it cannot have second thoughts, because it does not construct a representation of the experienced sequence, and it never reconsiders earlier decisions. We take the occurrence of second thoughts to imply retrospection; that is, a reconsideration of earlier decisions about change points in the light of later observations. We take the requirement that a model produce second thoughts to be another strong constraint on viable models.

### A Bayesian Change-Point Model With Evidence-Triggered Updating

We propose that the perception of Bernoulli probability is a by-product of the real-time construction of a compact encoding of the evolving sequence by means of change points, the points where the value of the hidden parameter is perceived to have changed. Our model has affinities with recent models of other perceptual processes that use Bayesian computation to construct an approximation to a minimum-description length representation (Feldman, 2009). Our model finds an approximation to the smallest sequence of change points that adequately describes the observed sequence, a description that is as simple as possible but not too simple.

Estimating the current probability would be easy if the subjects knew where the last change point was. Given the last change point, the estimate is simple Bayesian parameter estimation, using only the observations made since that change. We propose, therefore, a model in which subjects keep track of the change points, sometimes adding a new one, sometimes moving the most recent one a bit, and sometimes expunging the most recent one, as they see more data. Their current perception of probability is always based only on the data observed since the last change point in their representation of the sequence.

Figure 13 diagrams the flow of computation and Table 1 introduces our notation:  $p_g$  is the hidden parameter of the Bernoulli process and  $\hat{p}_g$  is the corresponding subjective probability (i.e., the subject's current estimate of  $p_g$ );  $p_c$  is the probability of a change

in  $p_g$  after any given trial (fixed at .005) and  $\hat{p}_c$  is the subject's perception of that change probability.

Beginning with a slider setting,  $\hat{p}_g$ , the computation proceeds in two steps. First, it decides whether the current estimate needs changing; if so, it then decides on why the estimate needs changing. Does it need to be changed because it was a misestimate based on too small a sample? Or does it need to be changed because there has been a (further) change in the hidden parameter? Or, does the problem arise because the most recently decided on change point should be taken back in the light of observations made subsequent to that most recent decision? This last possibility gives rise to second thoughts.

**First stage: Is it broke?** The first decision—whether there is a problem with the current percept—is based on the extent to which the number of green circles observed since the last change point deviates from the number predicted by the current percept. The expected number is  $n\hat{p}_g$ , where  $n$  is the total number of circles seen since the last change point (or since the start of the session), with standard deviation  $\sqrt{n\hat{p}_g(1-\hat{p}_g)}$ . If the actual number is too far from  $n\hat{p}_g$ , with “too far” measured relative to the standard deviation, the current percept is no longer consistent with what has been observed since that estimate was made. The process by which the first stage decides whether the current estimate is no longer viable is equivalent to a conventional null hypothesis significance test, with, of course, a decision criterion (alpha level). This first stage implements the principle that “if it ain't broke, don't fix it.” That principle explains subjects' intermittent slider adjustments: Most of the time their percept isn't broke, so they don't fix it.

The decision on whether the current estimate is broke must be based on a measure of the strength of the evidence that there is a problem. We denote this measure by  $E$ . We suggest that it is the product of the number of observations made since the most recently estimated change point ( $n_{>c}$ ) and the extent to which the proportion observed since that change point ( $p_o$ ) diverges from the  $\hat{p}_g$  made when that change point was added. As shown in Appendix B, the extent of the divergence is measured by the Kullback–Leibler divergence,

$$E = n_{>c} D(p_o \parallel \hat{p}_g), \quad (1)$$

where

$$D(p_o \parallel \hat{p}_g) = p_o \log \frac{p_o}{\hat{p}_g} + (1 - p_o) \log \frac{1 - p_o}{1 - \hat{p}_g} \quad (2)$$

and logs are natural logarithms. The Kullback–Leibler divergence measures the directed distance from one distribution to another; in this case, from Bernoulli ( $p_o$ ) to Bernoulli ( $\hat{p}_g$ ). It is the information-theoretic measure of how far one distribution diverges from another. It is the rate at which the odds that the observed distribution differs from the assumed distribution increases as one samples from the observed distribution. Thus, the product of the divergence and the number of draws measures the strength of the evidence for a discrepancy. An appeal of this measure is its computational simplicity. Because  $\hat{p}_g$  does not change from one observation to the next, the only quantity that is updated is  $p_o$ . Its new value depends only on  $n$  and  $n_g$ . On any observation,  $n$  increases by 1 and  $n_g$  by 1 or 0.

On the great majority of trials,  $E$  is less than any plausible decision criterion, which means that the estimate of the hidden

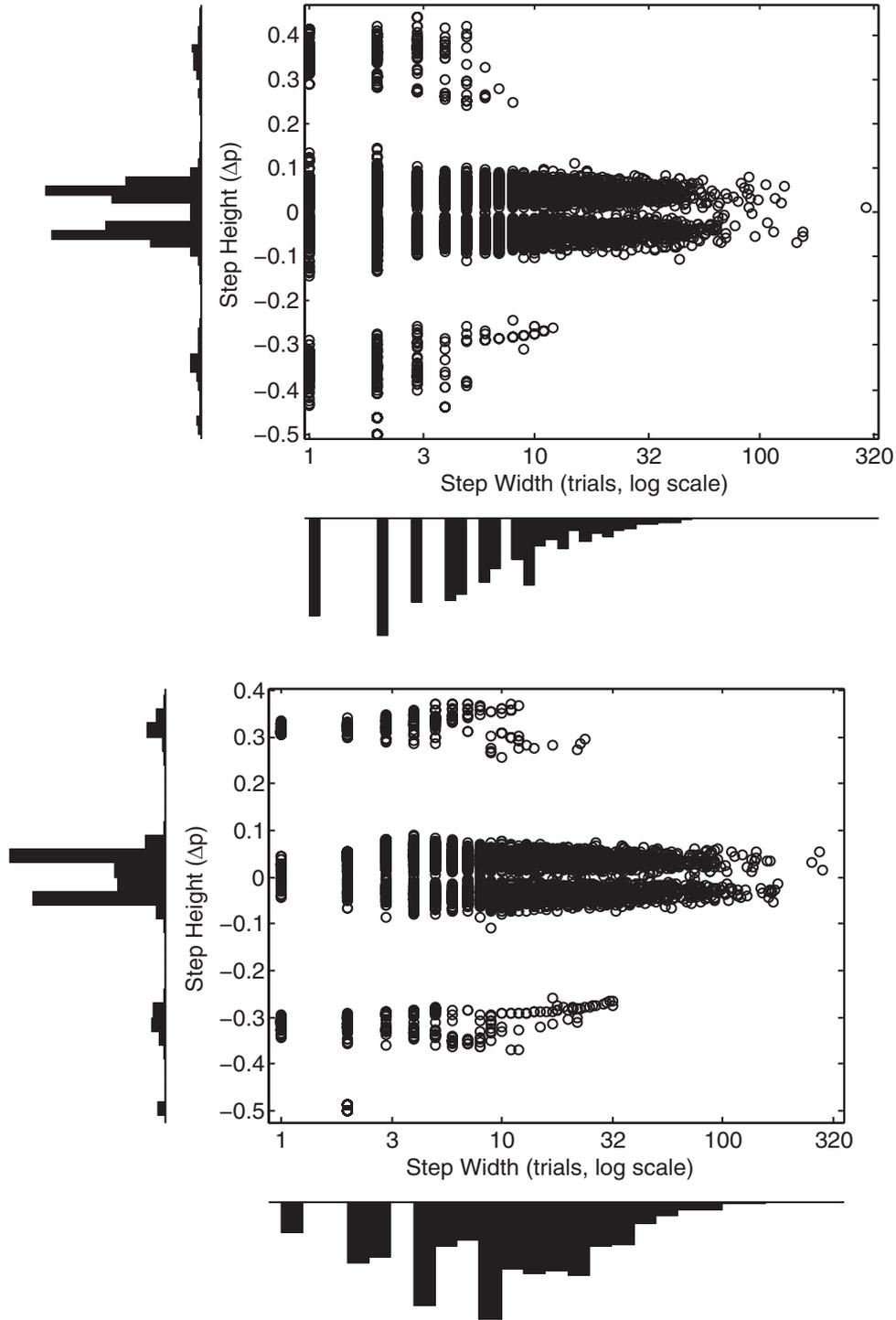


Figure 12. Joint and marginal distributions of step widths and step heights produced by a complex 5-parameter, 2-kernel, running-average model with a stochastic difference threshold on the output. In the upper panel, the rapid decay rate is relatively fast ( $\alpha = .15$ ), whereas in the lower it is slow ( $\alpha = .05$ ). In both cases the slow decay rate is four times smaller ( $\alpha = .0375$  and  $0.0125$ , respectively). The threshold difference required to switch the underlying estimate from the slow-decay estimate to the fast-decay estimate,  $\Delta_c$ , was  $0.3$ . The difference thresholds on the output were drawn observation by observation from a normal distribution with a mean,  $\mu_\tau$ , of  $.05$  and standard deviation,  $\sigma_\tau = cv_\tau \mu_\tau = .33 \times .05 = .0165$ . The online supplemental materials contain a custom MATLAB function, which may be used to run the model with parameter values of one's choosing.

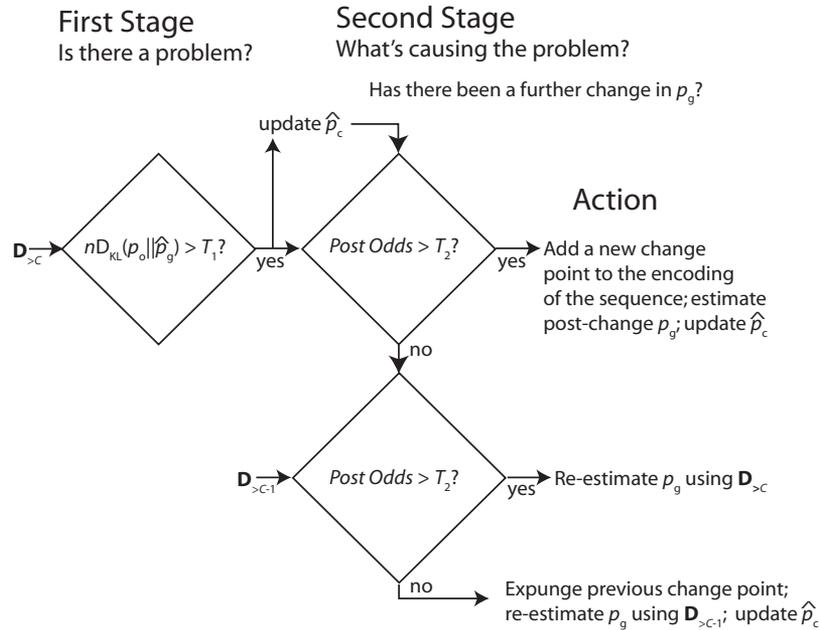


Figure 13. Flow of the computations in the proposed model.  $\mathbf{D}_{>c}$  is the sequence observed since the most recent change point (or the start of the session);  $p_o = n_g/n$  is the observed probability since the last change point (number of green circles divided by total number of circles);  $\hat{p}_g$  is the current estimate of  $p_g$ , which is the current true probability of a green circle. The posterior odds (PostOdds) are those in favor of an encoding of a binary sequence  $\mathbf{D}$  that inserts a change point somewhere in it, as opposed to an encoding that does not (the no-change encoding). The input to the lower decision diamond,  $\mathbf{D}_{>c-1}$ , is the sequence observed since the penultimate change point.  $T_1$  and  $T_2$  are the free parameters of the model (the decision thresholds). See Figure 14 for an illustration of the compactness and precision of the encodings generated by the model.

parameter is not broken. In that case, there is no further computation. This is one of the features that make the model computationally efficient. Another efficient feature of this measure is that it has a fixed distribution. We show in Appendix B that  $E$  is distributed as  $\Gamma(.5,1)$ . Thus, the first-stage decision (“Is it broke?”) employs a fixed decision criterion, unlike the decision criteria in most null hypothesis significance tests, which depend on the degrees of freedom. The decision criterion on  $E$  is  $T_1$ , the first of the two free parameters in our model.

**Second stage: Why is it broke?** When a new estimate for  $\hat{p}_g$  becomes necessary—when  $\hat{p}_g$  is broke—choices must be made in determining what the new percept should be. The appropriate fix depends on the answer to the question, What’s wrong? The problem may arise from small-sample estimation error. Initial estimates and estimates made when a change has just been detected are based on small numbers of observations, so they are often somewhat off. However, the problem may be that there has been a change in the true probability of a green circle. Or, the last change point may have been erroneous and should be expunged.

Multiple changes since the last change point are also a possibility, but, for reasons we discuss below, we assume they can be ignored. That leaves only two possibilities beyond small-sample estimation error: one additional change point or one expunged change point. In deciding between these possibilities, the process decides between two encodings of the sequence since the last change point: a simple one that uses only a single  $\hat{p}_g$  or a more complex one that uses a change point and two  $\hat{p}_g$ s, one before and one after the change point.

The decision between encodings with and without a change point is based on the posterior odds favoring one change somewhere in that sequence being considered versus no change. The posterior odds are the Bayes factor in favor of a change multiplied by the prior odds of a change,

$$PostOdds = BF \times PriorOdds = \frac{MML(M_1 | \mathbf{D}, \pi(\theta | M_1))}{MML(M_0 | \mathbf{D}, \pi(\theta | M_0))} \times \frac{p(M_1)}{p(M_0)},$$

where MML denotes the marginal (that is, integrated) likelihood of a model for the data (see Equations 3 and 4 below).

The Bayes factor is the relative likelihood of the two contrasting encodings (models), given the data,  $\mathbf{D}$ , and the prior distributions on the parameters,  $\theta$ , which for us consists of three quantities:  $\theta = (p_g^b, p_g^a, t)$ , where  $p_g^b$  and  $p_g^a$  are, respectively, the probability of a green circle before the change and the probability after the change, and  $t$  is the trial after which the change is estimated to have occurred. The one-change description uses all three parameters. The no-change description uses only one of them because under that description  $p_g^b = p_g^a$  and  $t = n$ ; that is, the change occurs notionally after the end of the sequence.

The Bayes factor depends on the distribution of the green circles within the sequence: Do they cluster toward one end or the other, or are they more or less evenly sprinkled throughout the sequence? The prior odds do not depend on this pattern; they depend only on the length of the sequence and on  $\hat{p}_c$ , the current estimate of how often changes occur.

Table 1  
Notation

Symbol or expression	Refers to
$p_g$	The true proportion of the green circles in the box at any one time; the true value of the hidden parameter of the Bernoulli process.
$\hat{p}_g$	The subject's current estimate of $p_g$ , as taken from the subject's current slider setting.
$p_c$	The probability that $p_g$ changes after any given trial. In our experiment, $p_c = .005$ .
$p(p_c)$	The subject's posterior probability distribution on $p_c$ .
$\hat{p}_c$	The expectation of $p(p_c)$ ; the subject's current estimate of the probability of a change in $p_g$ .
$\mathbf{D}$	A sequence of red and green circles (a binary data vector).
$C$	The index number of the latest change point in the subject's encoding of $\mathbf{D}$ ; $t_c$ is the trial number of the last change point.
$\mathbf{D}_{>C}$	The data (circles observed) since the latest change point.
$\mathbf{D}_{>C-1}$	The data since the penultimate change point.
$n_g, n_r,$ and $n$	Number of green circles, number of red circles, and number of trials, respectively, in a subsequence of trials.
$p_o = n_g/n$	The observed proportion of green circles in a subsequence.
$p(p_g)$	The Bayes-optimal real-time posterior distribution on $p_g$ after any given trial.
$\bar{p}_g$	The expectation of $p(p_g)$ . We use this as the trial-by-trial "observed" $p_g$ , the standard against which we measure the accuracy and precision of $\hat{p}_g$ , the subjects' trial-by-trial slider settings.
$D_{KL}(p_o \parallel \hat{p}_g)$	The Kullback–Leibler divergence of $\hat{p}_g$ (the slider setting) from $p_o$ ; a measure of how far the subject's perception diverges from the observed proportion. See Equation 2.
$n_d$	The expected number of trials an observer would need to find evidence of error. The smaller the error in the subject's estimate, the more trials are on average required to reveal it. $n_d$ is inversely proportional to the Kullback–Leibler divergence (the magnitude of the error); see Equation 5.
$L(\theta \mid \mathbf{D}, M)$	The likelihood function in Bayes's rule. $M$ is a model of the generative process; $\theta$ is the set of parameters whose values are estimated from the data when using that model to describe the data. The likelihood function specifies the uncertainty about these values given the information in the data.
$\pi(\theta \mid M)$	The prior distribution on $\theta$ in Bayes's rule. This distribution specifies the uncertainty about the values of these parameters before looking at the data.
$MML(M \mid \mathbf{D}, \pi(\theta \mid M))$	Bayesian marginal model likelihood of an encoding or description of the data; the integral of the product of the likelihood function and the prior distribution. The ratio of two marginal model likelihoods is the Bayes factor.
$\alpha_g$ and $\beta_g$	The (hyper)parameters of the subject's beta distribution prior on $p_g$ .
$\alpha_c$ and $\beta_c$	The (hyper)parameters of the subject's beta distribution prior on $p_c$ .
$Beta(\alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}$	The beta distribution, which is the conjugate prior for the Bernoulli distribution. $B$ is the beta function (not to be confused with the beta distribution).
$T_1$ and $T_2$	Respectively, the threshold (decision criterion) on $E$ (the evidence of a problem) and the threshold on posterior odds of a change.

$$\frac{p(M_1)}{p(M_0)} = \frac{n\hat{p}_c}{1 - \hat{p}_c}$$

The one-change description is more complex (lengthier) but also more flexible; it allows two probabilities of a green circle, one before and one after the change point, and it allows the change to happen at any point in the sequence. The no-change description is less complex (shorter) but less flexible; it allows only one probability to describe the whole sequence (that is,  $p_g^b = p_g^a$ ) and change point (that is,  $t = n$ ).

The maximally likely one-change description is always better than the best no-change description. However, the inherent ability of a more complex and flexible model to represent data more minutely than a less complex and less flexible one is offset by the fact that  $\pi(\theta \mid M_1)$ , the prior probability distribution of the one-change model, is more diffuse than  $\pi(\theta \mid M_0)$ , the prior for the no-change model: The former occupies a three-dimensional space, whereas the latter occupies only a one-dimensional space; both have, of course, a unit mass of total prior probability distributed in the space they occupy. The Bayes factor adjudicates the trade-off between the descriptive adequacy and the complexity of the competing descriptions; that is, between the minuteness of a description and its lengthiness. It implements Einstein's razor, favoring

the description that is as simple as possible but no simpler (Gallistel, 2009; MacKay, 2003).

The prior odds take into the account a different consideration: When the sequence is short and the probability of change is low, the sequence probably does not contain a change, regardless of what the pattern of green circles within it may suggest. When, on the other hand, the sequence is much longer than the expected interval between changes ( $1/\hat{p}_c$ ), the a priori probability that there has been a change is high, even if the pattern of green circles within the sequence gives only weak support for this description. To correctly estimate the odds that there has been a change, the computation must take account of both the Bayes factor and the prior odds.

The probability of a change after any given trial,  $p_c$ , is always small relative to its complement ( $1-p_c$ ); therefore, the odds of a second change within the rather short sequences of trials required for an earlier change to become evident are low. That is why the process does not have to decide between one-change descriptions and descriptions with more changes. When the changes are small, they are hard to detect—it takes many trials—so there may well be other changes, but then the descriptive adequacy of the simpler descriptions suffers very little by neglecting those changes.

When the posterior odds in favor of a change exceed a decision criterion ( $T_2$  in Figure 13), the perceptual process adds a change point at the most likely point in the sequence seen since the last change. In doing so, it lengthens its change-point description of the overall sequence.

When the algorithm decides that an additional change point is not needed to describe the sequence since the last change point, it then decides whether the previous change point should be expunged. To do this, it provisionally expunges the last change point, then proceeds to decide, as above, whether the sequence since the penultimate change point is better described with or without a change point. When the decision is in favor of a change point, the expunged change point is restored but not always at precisely the same point in the sequence; otherwise, it is not restored, and what was the penultimate change point becomes the most recent one. That is, the model has a second thought.

When all has been decided, the new percept is the relative frequency of the green circles since what has now been decided to be the latest change point, slightly modified by the prior distribution on  $p_g$ :  $\hat{p}_g = (n_g + \alpha_g)/(n + \alpha_g + \beta_g)$ . Because changes are often detected quickly, when a new change point is added,  $n$  is generally small. Consequently, those new estimates are often inaccurate. When the last change point has been expunged, on the other hand,  $n$  is generally large, so these percepts are generally accurate. The construction of the encoding of the overall sequence is iterative (that is, tail recursive): The new estimate is now checked after each new circle to see if it is broke; when it is broke, fixing it proceeds as above.

Figure 14 illustrates the descriptive adequacy and the compactness of the encodings produced by this model. There, we plot the cumulative records of the green circles seen by the first six subjects in their first session (thin black solid, somewhat wiggly lines). On these records, we superimpose the change points by which the model represents each sequence. These change points are connected by thick dashed gray smooth lines, which are the cumulative records recoverable from the encoding. The point slope of a cumulative record is the relative frequency (probability) of green circles in the vicinity of that trial. The slopes of the thick gray dashed lines are everywhere extremely close to the slopes of the thin black wiggly lines, indicating that the representations are descriptively adequate, yet the change-point encodings in Figure 14 have only 3–6 change points per sequence of 1,000 stimuli. (The circle at the origin of a plot is not counted, because its coordinates are always 0,0.) The coordinates of a change point form a two-dimensional vector; thus, 8–12 numbers suffice to describe (summarize) a binary sequence of 1,000 stimuli, a reduction by two orders of magnitude in the number of symbols employed.

**Computational efficiency.** On the relatively rare occasions when the algorithm enters the second stage, it first updates its estimate of the probability of a change. We assume a conjugate (Beta) prior on  $p_c$ , so the updating computation is trivial,

$$\hat{p}_c = \frac{n_c + \alpha_c}{n + \alpha_c + \beta_c},$$

where

$n_c$  = number of changes so far perceived, across all sessions.  
 $n$  = the total number of trials, across all sessions.

$\alpha_c$  and  $\beta_c$  = the parameters of the Beta prior on  $p_c$ .

For most modeling purposes,  $\alpha_c$  and  $\beta_c$  can be set to .5 (the Jeffreys prior). The estimated probability of a change,  $\hat{p}_c$ , determines the prior odds of a change in a sequence of any given length. The prior odds of a change play an important role in the decision as to the proper explanation for the problem, because the posterior odds in favor of a change are the product of the Bayes factor and the prior odds.

The Bayes factor is the ratio of the marginal likelihoods of two competing descriptions of the sequence observed since the last change. We use a conjugate prior, Beta( $\alpha, \beta$ ), on the possible values of the probability before and after a putative change. With this prior, the computation of parameter estimates and marginal likelihoods is simple and efficient. The marginal model likelihood (MML) of the one-change encoding is

$$\text{MML}(M_1 | \mathbf{D}, \pi(\theta | M_1)) = \iint dp_g^b dp_g^a L(p_g^b, p_g^a, t | \mathbf{D}) \pi_1(p_g^b, p_g^a, t), \quad (3)$$

where  $L(p_g^b, p_g^a, t | \mathbf{D})$ , the likelihood of the parameter values given the before and after change probabilities, is given by

$$L(p_g^b, p_g^a, t | \mathbf{D}) = (p_g^b)^{n_g^{\leq t}} (1 - p_g^b)^{n_r^{\leq t}} (p_g^a)^{n_g^{\geq t}} (1 - p_g^a)^{n_r^{\geq t}}.$$

Here,  $t$  is the trial index,  $n_g^{\leq t}$  and  $n_r^{\leq t}$  are the numbers of green and red circles observed on Trials 1 to  $t$ , and  $n_g^{\geq t}$  and  $n_r^{\geq t}$  are the numbers of green and red circles observed after  $t$ . The prior,  $\pi_1(p_g^b, p_g^a, t)$ , is a product of beta distributions, plus a factor of  $1/n$  indicating that there is an equal probability of a change on every trial,

$$\pi_1(p_g^b, p_g^a, t) = \frac{(p_g^b)^{\alpha_g - 1} (1 - p_g^b)^{\beta_g - 1} (p_g^a)^{\alpha_g - 1} (1 - p_g^a)^{\beta_g - 1}}{B(\alpha_g, \beta_g) B(\alpha_g, \beta_g) n},$$

where  $B(\alpha, \beta)$ , the beta function, is the usual normalizer for the beta distribution. The integrals in Equation 3 are straightforward, and we have

$$\begin{aligned} \text{MML}(M_1 | \mathbf{D}, \pi(\theta | M_1)) \\ = \frac{1}{n} \sum_{t=1}^n \frac{B(n_g^{\leq t} + \alpha_g, n_r^{\leq t} + \beta_g) B(n_g^{\geq t} + \alpha_g, n_r^{\geq t} + \beta_g)}{B(\alpha_g, \beta_g) B(\alpha_g, \beta_g)}. \end{aligned}$$

For the no-change model,  $p_g^b = p_g^a$  and  $t = n$ , in which case the marginal likelihood simplifies to

$$\text{MML}(M_0 | \mathbf{D}, \pi(\theta | M_0)) = \int dp L(p | \mathbf{D}) \pi_0(p), \quad (4)$$

where

$$L(p | \mathbf{D}) = p^{n_g} (1 - p)^{n_r}$$

and

$$\pi_0(p) = \frac{p^{\alpha_g - 1} (1 - p)^{\beta_g - 1}}{B(\alpha_g, \beta_g)}.$$

The integral in Equation 4 is also straightforward, and so, for the no change model,

$$\text{MML}(M_0 | \mathbf{D}, \pi(\theta | M_0)) = \frac{B(n_g + \alpha_g) B(n_r + \beta_g)}{B(\alpha_g, \beta_g)}.$$

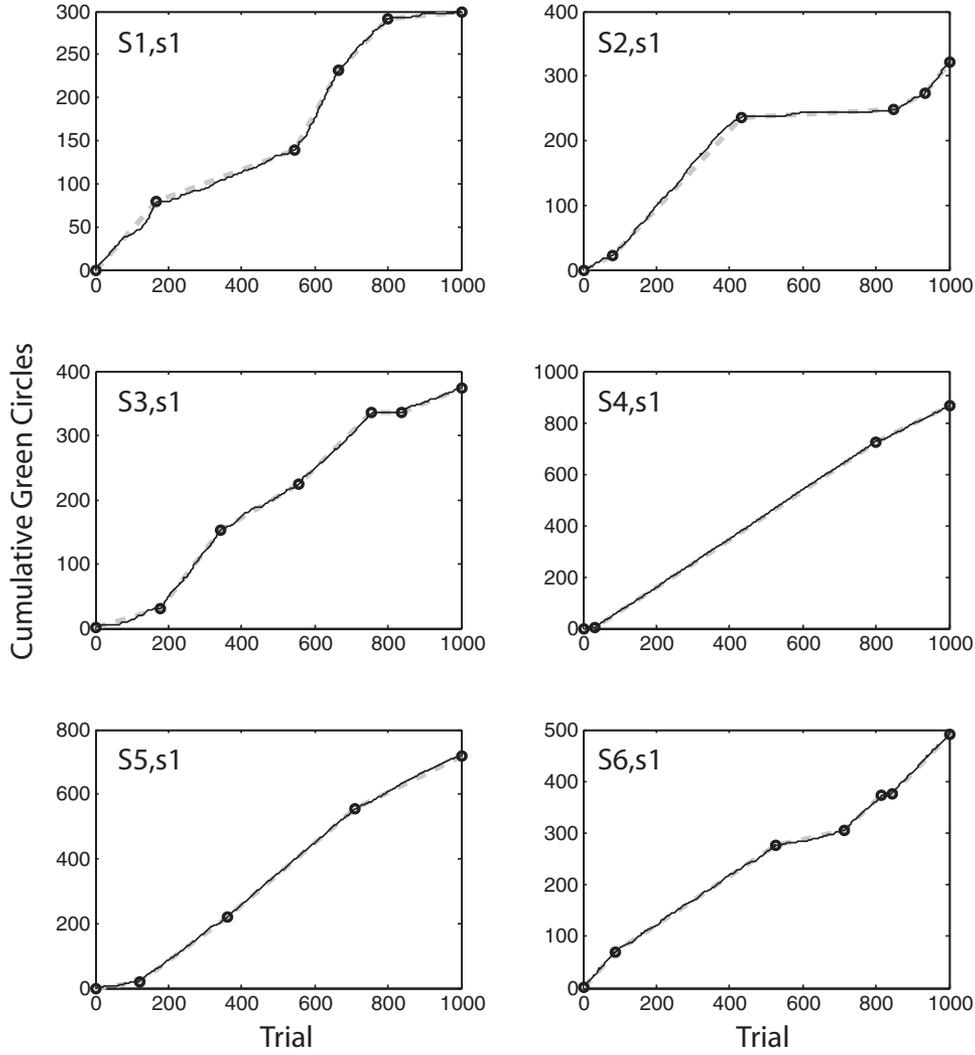


Figure 14. The model’s change-point representation of the stimulus sequences seen during their first session by the first six subjects (circles connected by thick dashed gray lines) superimposed on the cumulative records of the green circles seen during those sequences (thin solid black lines). The point slopes on these cumulative records are the “observed”  $p_g$ s; that is, the relative frequency of the green circles in the vicinity of a given trial. Note how closely the observed sequences may be represented by a few change points. These representations were computed with the optimal decision criterion on the posterior odds of a change.

The posterior odds are, then, given by

$$PostOdds = \frac{n\hat{p}_c \text{ MML}(M_1 | \mathbf{D}, \pi(\theta|M_1))}{1 - \hat{p}_c \text{ MML}(M_0 | \mathbf{D}, \pi(\theta|M_0))}$$

When the posterior odds exceed a decision criterion, the value of  $t$  at which the likelihood function peaks is taken to be the trial after which  $p_g$  changed. This new change point is added to the encoding of the sequence, and the data after it are used to compute the new  $\hat{p}_g$ .

All the computations are in closed form; no integration, numerical or stochastic, is required. The model is fully implemented by a custom MATLAB function, ChangePointModel.m, which is provided in the supplemental materials.

**Estimating parameter values.** Our model has only two free parameters:  $T_1$ , the decision criterion on the evidence for a problem with the current percept, and  $T_2$ , the decision criterion on the posterior odds in favor of a change description. The model generates simulated results for many different aspects of subjects’ performance: number of steps, step widths, step heights, hit and false alarm rates for change detection, detection latencies, number of second thoughts, and second thought latencies, among others. It is not possible to search the two-dimensional space of model parameters for the values that do the best job, because the output space (the dependent variables) is for the most part mutually incommensurable. There is no obvious way to weight misestimates of the number of steps produced relative to misestimates of the

false alarm rate or misestimates of the distribution of change-detection latencies.

Given this, we chose to adjust  $T_1$  and  $T_2$  to match as closely as possible the number of slider steps the subjects made and their hit and false alarm rate. We did this in two steps. The first was to choose  $T_1$ . To do that, for each subject-specific set of 10 stimulus sequences (each consisting of 1000 trials), we found the lowest value for  $T_1$  such that the model produced very nearly the same number of steps as did the subject (within 5%) at five values of  $T_2$ : 1, 2, 4, 8, and 16. This was not possible for one subject; in that case, we used the value of  $T_1$  that, when  $T_2 = 1$ , produced the same number of (simulated) steps as did the subject. The second step was to choose  $T_2$ . We chose for each subject two values for it to match, as closely as possible, the subjects' actual hit and false alarm rates. For the one errant subject, we set  $T_2$  to 1. Referring back to Figure 8, one sees that with  $T_1$  fixed at values that give close approximations to each subject's number of slider steps, there are in most cases values of  $T_2$  that give simulated hit rates (green forms in Figure 8) and false alarm rates (red forms in Figure 8) reasonably close to those produced by each subject.

We then asked whether the joint distribution of step widths and step heights that the model generated with the same parameters was a reasonable approximation to the subjects' distribution. As may be seen in Figure 15 (on the right), the answer is again at least a qualified yes. There is, though, room for improvement. The model's marginal distribution of step heights is more sharply peaked than the subjects'; that is, the model makes many more very small adjustments. The model's distribution of step widths declines almost monotonically, whereas the subjects' peaks at a width of a little less than 10. The size of the model's adjustments tapers off as step widths get longer, whereas the size of the subjects' adjustments does not. In any event, the approximation to the joint distribution is much better than the best we could obtain from a complex running-average model with three more free parameters.

Finally, we asked whether, within the same range of parameter combinations, there exists a combination that produces reasonable approximations to the distributions of change-detection latencies and second thought latencies. To answer that we set the parameters at  $T_1 = .82$  and  $T_2 = 2$  (approximately the centroid of the range of values used for the results in Figure 8) to obtain the model results shown in Figure 7 (change-detection latencies) and Figure 9 (second thought latencies). As one sees in those two figures, here, too, the distributions of detection latencies and second-thought latencies generated by the model are reasonable approximations to the distributions generated by the subjects.

**Choice of priors.** In discussing the results so far, we have not treated the parameters of the prior distributions as free parameters. We have used so-called objective parameters for our beta distribution priors on  $p_g$  and  $p_c$ . The parameter values generally considered to make the beta distribution an objective prior are Beta(0,0), the Haldane prior; Beta(.5, .5), the Jeffreys prior; and Beta(1,1), the Bayes–Laplace prior.

The Haldane prior does not bias the posterior distribution at all, which means, for example, that if the first two observations are both 1 (“heads” or “successes”), then the posterior distribution, Beta(2,0) asserts that 1 is the only value for  $p$  that has nonzero probability. Thus, the expectation at that point is that the sequence of 1s will continue uninterrupted indefinitely. Most users find this

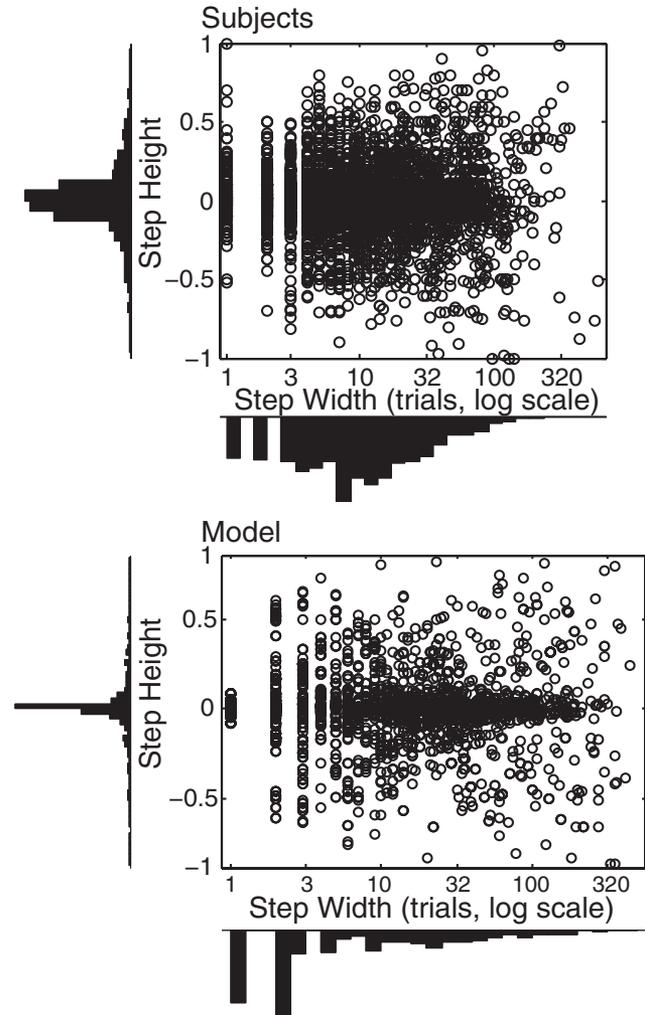


Figure 15. Joint and marginal distributions of step widths and step heights from subjects (top panel, reproduced from Figure 11) and from the change-point model (bottom panel) when the decision thresholds have been adjusted for each subject to simulate that subject's approximate number of steps and approximate false alarm and hit rates in detecting changes.

unreasonable, and so do we. Also, with this prior, there is no a priori estimate of the average probability of a green circle, because Beta (0,0) is not normalizable.

The Jeffreys prior and the Bayes–Laplace prior both bias the estimate of  $p$ , the Jeffreys prior less so than the Bayes–Laplace prior. They deliver the same a priori estimate of the average probability of a green circle (namely, .5), which is in accord with our subjects' expressed estimate (we asked subjects to set the slider even before they saw the first circle, and they all set it close to .5).

The Bayes–Laplace prior accords with one's intuition about a neutral prior in that it is uniform; it asserts that all values of  $p$  are equally probable a priori. The Jeffreys prior strongly violates this intuition in that it is cup shaped; it treats values of  $p$  at the extremes as much more probable a priori than values in the middle. However, it does have two advantages: It is less informative than the

Bayes–Laplace prior, and, unlike the Haldane prior, it is normalizable.

As already noted, the Haldane prior gives what is generally regarded as an unacceptable posterior distribution when the first two or three observations are the same, as they often are. It is interesting to note what the alternative objective priors do in this regard. The estimated number of trials to a failure using the Bayes–Laplace prior after  $n$  successive initial successes is  $n + 2$ , so the sequence already experienced is always within 2 of the expectation. Using the Jeffreys prior, the expected number of trials to a failure is  $2n + 2$ , so the already experienced sequence is always slightly less than half the expectation.

In the modeling so far described, we used the Jeffreys prior on  $p_c$ , because we think it the most reasonable objective prior, and the Bayes–Laplace prior on  $p_g$ , because the first author was under the mistaken impression that the closed-form solution for the posterior marginal likelihoods was not valid for the Jeffreys prior. When this impression was corrected, he ran the model again with a Jeffreys prior on both  $p_g$  and  $p_c$ . The results, although not identical, were so similar that we have not changed the figures.

### The Perception of Volatility and the Prior on $p_c$

A result we have not so far mentioned is that 9 of the 10 subjects showed a higher rate of change detection in the first two sessions than in the last two sessions (see Figure 16, left panel). This suggests that subjects began the experiment with some belief about how frequently they would encounter changes, that these initial beliefs were on the high side, and that subjects revised them downward toward the true value as they gained experience. This conclusion accords with the conclusion from other recent experiments and with models cited in our introduction (e.g., Behrens et al., 2007), which have assumed that subjects estimate volatility and that their estimates of volatility affect their rapidity of their response to changes (something that Robinson also observed).

We capture this phenomenon with a nonobjective prior on  $p_c$ , which of course introduces two additional free parameters; namely, the parameters of the beta distribution on  $p_c$ . The choice of these parameters reflects two assumed degrees of freedom in a

subject’s initial (not necessarily conscious) belief about the probability of a change in  $p$ . The ratio  $\alpha_c/(\alpha_c + \beta_c)$  captures what the subject thinks that probability is likely to be, and  $\alpha_c\beta_c/[(\alpha_c + \beta_c)^2(\alpha_c\beta_c + 1)]$  is the variance of that belief. As may be seen from the right panel in Figure 16, giving the model such a prior does indeed produce a downward drift similar to that observed in the subjects’ data.

In summary, our model implements the kind of model Robinson believed was necessary—a model based on a sequence of decisions—using (for all but one purpose) only two free parameters, the decision criteria,  $T_1$  and  $T_2$ . There are no learning-rate parameters; that is, no parameters that determine the range of past data over which an average is taken or upon which a Bayesian estimate is based. And, except when we attempt to account for subjects’ slowly decreasing frequency of change detection, there are no volatility parameters, parameters that affect estimates of the frequency of change. When we use the objective Jeffreys prior on  $p_c$ , (or the Bayes–Laplace prior), the model captures the effect of varying objective volatility without positing a free parameter for that purpose. It does so because it estimates  $p_c$  from the data (with a slight bias from an objective prior). Its estimate ( $\hat{p}_c$ ) determines the prior odds of a change in a sequence of any given length, and varying the prior odds varies the sensitivity to within-sequence evidence of a change. Thus, subjects encountering high volatility will be more likely to detect changes in a sequence of any given length than will subjects encountering low volatility. We adjust the parameters of the prior on  $p_c$  only when we attempt to capture the slow adjustment to observed volatility actually shown by our subjects. For that purpose, the Jeffreys prior and the Bayes–Laplace prior are too weak; they are overcome by the data too fast.

### The Precision of Subjects’ Estimates

We have deferred discussion of this seemingly purely empirical question because discussing it required two theoretical developments, an algorithm for computing the ideal observer’s estimate of probability on a trial-by-trial basis and an appropriate measure of the error (i.e., the difference between a subject’s slider setting on a given trial and the estimate of the ideal observer of the sequence seen by the subject up to that trial). By “error” we do not mean how far the slider setting was from the true probability of a green circle,  $p_g$ . Instead, we mean how far it was from what we call the observed  $p_g$ , denoted  $\bar{p}_g$ , which is the posterior mean of a Bayes-optimal probability tracking algorithm that is closely related to that published by Adams and Mackay (2006). The mathematical details are in Appendix C. The importance of using a measure based on what subjects have actually observed, as opposed to the true but hidden  $p_g$ , which they have not observed, is shown in Figure 17. There we plot, for each subject, the root mean square error between the slider setting and the true  $p_g$  and that between the slider setting and the “observed” probability, as defined by the Bayes-optimal probability tracking algorithm. Robinson’s measure of precision corresponded to the former (subject’s  $p_g$  vs. true  $p_g$ ), but in every case, the latter (subject’s  $p_g$  vs. observed  $p_g$ ) is smaller, as one would expect.

For the second question, the measure of precision, we use the inverse of the Kullback–Leibler divergence from the Bernoulli distribution with  $p = \bar{p}_g$  to the Bernoulli distribution with  $p = \hat{p}_g$  (the slider setting). The inverse of the Kullback–Leibler diver-

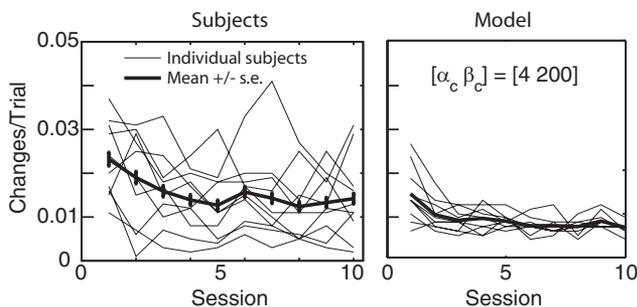


Figure 16. Downward drift in the frequency of change detection. Left panel: Individual subjects (thin lines) and the overall session-by-session mean across the subjects (heavy line). Right panel: Model with prior parameters of the beta distribution on  $p_c$  set to  $\alpha_c = 4$  and  $\beta_c = 200$ , corresponding to  $\bar{p}_c = 4/204$  [approx] .02 (4 times higher than the true  $p_c$ ). Thin lines and heavy lines are the same as in the left panel. s.e. = standard error.

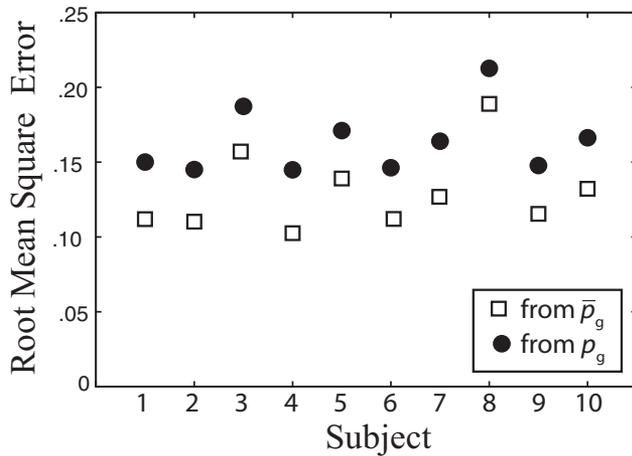


Figure 17. The root mean square error between the subjects' slider settings and the ideal observer's estimates ( $\hat{p}_g$ , open squares), and between the slider settings and true values ( $p_g$ , filled circles). As expected, subjects' percepts are closer to the observed value than to the hidden true value.

gence is the expected number of trials an ideal observer would take to detect that the slider setting is wrong (taking the ideal observer to be ground truth; see Appendix B). Given a conventional decision criterion (alpha level) of .05, the expected number of trials to discover the error in a slider setting is

$$n_d = \frac{1.92}{D_{\text{KL}}(\bar{p}_g \parallel \hat{p}_g)}. \quad (5)$$

To address the question of whether  $n_d$  varied as a function of  $\bar{p}_g$ , we created five bins for the  $\bar{p}_g$  values—[.4–.6], [.3–.4 and .6–.7], [.2–.3 and .7–.8], [.1–.2 and .8–.9], and [0–.1 and .9–1]—and computed the cumulative distributions of  $n_d$  within each bin (across the 10 subjects). The five cumulative distributions of  $n_d$  are superposed in Figure 18. When  $\bar{p}_g$  is between 0.1 and 0.9 (all but the thinnest line), the distributions are virtually identical. For  $\bar{p}_g$  below 0.1 or above 0.9 (the thinnest line), the number of trials to disconfirmation is slightly higher, meaning subjects do slightly better. Again, this probably reflects the fact that subjects occasionally set the slider to 0 or 1 (see figure caption).

Our model's estimates are more precise than are even the best subject's, but we believe this is plausibly explained by the fact that the model computes with exact counts. Weber's law is known to apply to subjects' representations of the results of nonverbal counting (Cordes, Gallistel, Gelman, & Latham, 2007; Cordes, Gelman, Gallistel, & Whalen, 2001; Gallistel & Gelman, 2005). The imprecision in the brain's encoding of the counts must contribute to imprecision in its estimates of the ratios between counts.

## Discussion

### A Simple, Computationally Efficient Model

A model with two free parameters reproduces a wide range of results from a simple experiment in which subjects indicate, outcome by outcome, their perception of the current value of a stepwise nonstationary Bernoulli probability. Like the subjects, the

model delivers reasonably accurate estimates of probability over the full range of its possible values, not systematically over- or underestimating anywhere in this range. It explicitly detects changes (as opposed to simply adapting to them). It quickly detects substantial changes in probability, with a high hit rate and a low false-alarm rate. It often makes large discontinuous changes in the estimate of probability between one observation and the next, but it also often maintains a single, unvarying estimate for many trials. It explicitly distinguishes between, on the one hand, errors in the percept that become apparent as more data accumulate since the last change and, on the other hand, errors that arise because the hidden probability has changed again.

As experience accumulates, the model updates its estimate of the probability of a change in the Bernoulli probability. This estimate strongly affects the future probability of detecting a (putative) change. Thus, the model is sensitive to volatility (change probability), as are the subjects in our experiment and in other experiments (Behrens et al., 2007; Nassar et al., 2010, 2012; Robinson, 1964; Wilson et al., 2010). Finally, it has second thoughts about some previous change perceptions, as did our

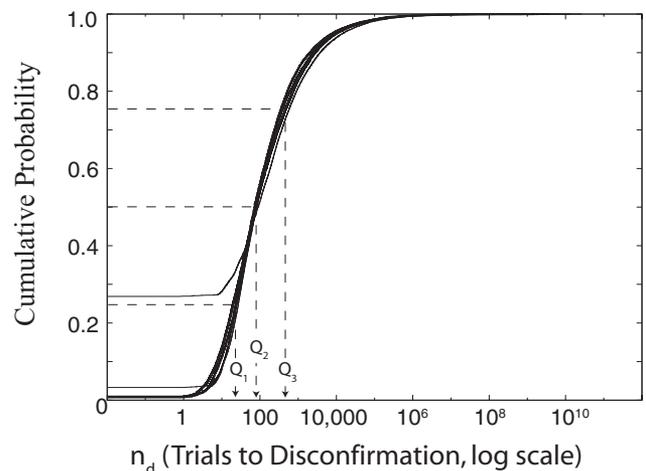


Figure 18. Cumulative distributions of the errors (deviations of the slider settings from the ideal observer's setting, the latter taken to be ground truth), as measured by trials to discovery. Dashed lines indicate the quartiles. The median trial-by-trial discrepancy between a subject's estimate and the estimate of the ideal real-time observer was such that it would require approximately 100 trials worth of data to detect the error ( $Q_2$  arrow). The distributions from different bins superpose, which means that the magnitudes of subjects' errors are independent of the objective probability, when error is measured in this way. The only notable failure of superposition is for the thin curve that meets the ordinate at about .27. This curve is for the data in the most extreme bin (0–.1 and .9–1). The flat bottom of this curve arises from cases in which the slider (the subject's estimate) was at 0 or 1 on trials where the ideal observer's estimate was not quite that extreme. In such cases the divergence of the subject's estimate from the observed value is infinite no matter how small the discrepancy in the two probabilities. The divergence is infinite because the first occurrence of an unexpected circle gives unbounded confidence that the subject's extreme estimate is wrong. Thus, in a finite number of trials one will attain infinite confidence that the estimate is wrong, so the rate at which confidence grows must itself be infinite, and the Kullback–Leibler divergence is the measure of that rate.

subjects. The latencies with which these second thoughts occur are approximately the same as those of the subjects. On all of these measures, the performance of the model is in qualitative and, to a first approximation, quantitative accord with the performance of subjects. Moreover, varying its two decision criteria within a plausible range of values captures much of the between-subject variation in these measures of performance.

The model is computationally efficient for two reasons. A simple first-stage computation that determines whether there is a problem with the current percept (current estimate of  $p$ ) is all that is required on most trials, because there seldom is a problem. When there is a problem, the model decides among three possible explanations for it, using basic, closed-form Bayesian model selection computations. There is no numerical integration and no need for particle filters or other stochastic approximations to Bayesian integrals.

### The Representation of Continuous Change

The question arises how well a discrete-change model of the kind we propose can represent a process of continuous change. Estes (1984) had subjects choose which of two options was likely to pay off under circumstances where the probability that one would pay off stayed constant at .5, while the probability that the other would pay off varied from 0 to 1 in accord with a sine function with a period of 80 trials. For half the subjects, after 4 cycles, the option with the heretofore variable probability took on a constant probability of .6. The subjects' probability of choosing the variable-probability option fluctuated with approximately the same period as the variable probability, as predicted by Estes' delta-rule updating model. However, on the fifth cycle, the half of the subjects whose variable-probability option no longer varied varied their choice probability in a pattern very similar to the pattern exhibited by the subjects for whom this option continued to vary. This was not what the delta-rule updating model did; it converged fairly rapidly on a steady probability estimate. Estes concluded that people not only track probabilities; they construct a model of how those probabilities are likely to change in the future. This conclusion is in line with our model's emphasis on the

explicit representation of the changes in the hidden stochastic parameter.

Estes's plots show smooth variation, but the smoothness may be a consequence of the averaging that he had to do. (His subjects did not overtly estimate the hidden parameter; they made binary choices that were presumably based on estimates of those probabilities or some function of them. Thus, to get trial-by-trial estimates of his subjects' hidden choice probabilities, Estes had to resort to averaging, across trials, across subjects, or both.) It will clearly be important to run our experiment with continuously varying probabilities to see if we can confirm Robinson's informal report that subjects showed the step-hold pattern of probability estimates (lever-dial positions) even when the probabilities they were tracking varied continuously.

The fact that there is no simple a priori form for the representation of continuous variation is relevant in thinking about how the brain might solve the problem. As Estes (1984) remarked, "There is little theory available to indicate what features or aspects of the temporally varying reward function might have been encoded in memory" (p. 264). An attraction of the change-point approach to the representation of variation in the hidden parameter is that it makes no commitment to the form of the variation, while laying the foundation for the extraction of parameters that characterize that variation. Figure 19 shows a simulation of the condition in which the true probability of a green circle varies as a sine function with a period of 80 trials (black curve). The simulated occurrences of green circles are plotted as asterisks at the top, and the simulated occurrences of red circles are plotted as asterisks at the bottom. The red "curve" is the trial-by-trial slider setting generated by our model when given this stimulus, with its decision thresholds set to those representative of the subjects' ( $nD_{KL}$  threshold = .82; threshold on posterior odds of a change = 1). The green "curve" is our model's retrospective representation of the stimulus, the unobserved representation that has benefited from second thoughts.

The change-point representation facilitates the recognition of the periodic structure of the variation. That recognition makes for a still more compact encoding of this sequence. That the changes

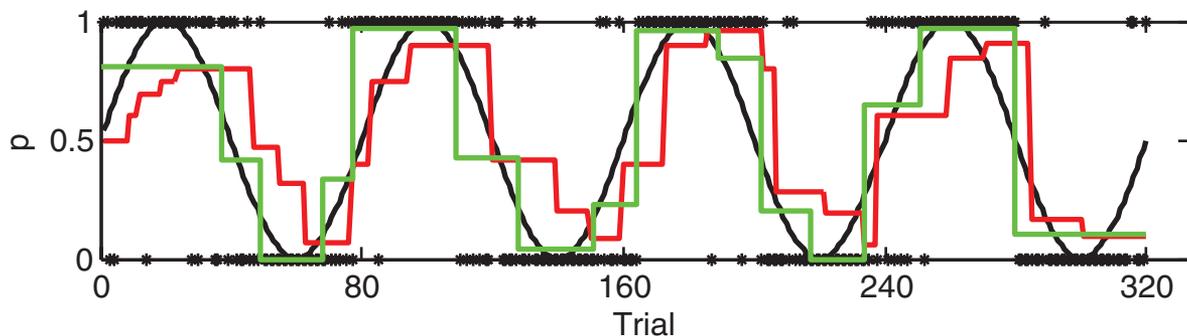


Figure 19. Simulation of sine-curve variation in the true Bernoulli probability as in Estes (1984). The black curve is the true hidden probability. The asterisks show the Bernoulli outcomes (green circle outcomes on top, red on bottom). The red step-curve is the slider setting that the model predicts a subject would produce given this continuously varying stimulus. The green step-curve is the model's prediction of the subject's retrospective representation of the sequence, the representation after second thoughts. It remains to be established whether subjects' slider settings will in fact look like the red curve.

occur periodically rather than randomly might be indicated by any of several simple tests on the intervals between the retrospectively revised change-point encoding (the one plotted in green in Figure 19). For example, the lower end of the 95% confidence interval on the shape parameter of the gamma distribution that best described the distribution of these intervals is 2.9. This lower limit is well above the value of 1, which is what the estimate of the shape parameter would approximate if the distribution of these change points were exponential. This simple test on the model's encoding of this sequence would conclude in favor of a nonrandom structure in the stimulus. Moreover, the average interval between reversals in the sign of the changes is 40.4, which is a good estimate of the half period of this stimulus.

In short, the change-point representation can represent fairly well stochastic parameters that change either discretely or continuously. In either case, it lays the foundation for the recognition and parameterization of whatever further structure there may be. The recognition of further structure makes possible a still more compact encoding. This still more compact encoding controls subsequent behavior insofar as subjects anticipate the continuation of that structure.

## Neuroscientific Implications

The linking of behavioral phenomena and cognitive mechanisms to neurobiological loci and neurobiological mechanisms depends on a computational theory (Carandini, 2012; Marr, 1982). We have here described a computational theory that differs in its fundamental assumptions from those that have been used to interpret recent neuroscientific findings.

Contemporary neuroscientific theorizing about the neurobiology of decision making under uncertainty is dominated by delta-rule updating models for the estimation of stochastic parameters (Behrens et al., 2007; Brown & Steyvers, 2009; Corrado et al., 2005; Courville, Daw, Gordon, & Touretzky, 2004; Courville, Daw, & Touretzky, 2006; Dayan & Daw, 2008; Glimcher, 2003a, 2003b, 2009; Montague et al., 1996; Nassar & Gold, 2013; Nassar et al., 2010, 2012; Seymour et al., 2004; Steyvers & Brown, 2006; Sugrue, Corrado, & Newsome, 2005; Wilson et al., 2010). They have a long history (Estes, 1957; Rescorla & Wagner, 1972). At their heart, there is a learning rate. The learning rate is the constant of proportionality in the delta rule; it relates the magnitude of the error to the magnitude of the adjustment in the estimate of the stochastic parameter.

Our model has no learning rate. It "learns" a new estimate only when it perceives a discrepancy between its current estimate and its recent experience. It bases its new estimate only on observations judged to have come from the stationary portion of its recent experience, the portion after the most recent change. More often than not, the new estimate is based on an entirely different set of observations from the set on which the old estimate was based. Thus, there often is no compromise struck between the testimonies of older and newer data.

Older theories with delta-rule updating assumed a constant learning rate. More recent work has attempted to measure the learning rate during periods of stability and instability (Behrens et al., 2007; Krugel et al., 2009; Mathys, Daunizeau, Friston, & Stephan, 2011; Nassar et al., 2010; Preuschoff & Bossaerts, 2007; Wilson et al., 2010; Yu & Dayan, 2005). This work has led to the

conclusion that the learning rate is itself modulated by a trial-by-trial updating process that estimates the volatility. During some stretches, subjects appear to learn slowly, which is good in that doing so yields more precise and stable estimates; whereas during other stretches, the same subjects appear to learn rapidly, which is good in that it enables rapid adaptation to a new state of the stochastic process.

The experimental results that inspired our model make it clear that if the underlying neurobiological mechanism or process really does operate in accord with the delta rule, then its learning rate must indeed be modulated by experience. Trial-by-trial records show long stretches in which the estimate changes not at all or only by very small amounts. These stretches can be explained by a delta-rule model only if the learning rate is assumed to be very small during those stretches. However, the same records show large step changes, changes from one trial to the next that span most of the possible range. These large steps can be explained by a delta-rule model only if the learning rate is assumed to be very high. In fact, empirical estimates of the rate have ranged as high as 1 (Nassar et al., 2010). When the learning rate is 1, the delta-rule estimate is based entirely on the most recent observation.

Our model calls into question the assumption that the underlying neurobiological mechanism or process operates in accord with the delta rule or, more generally, that it operates in accord with any rule in which the estimate after each observation is a compromise between the value suggested by that observation alone and the value suggested by a weighted average over preceding observations. Our model shows that a process operating in accord with quite different computational principles produces both the long stretches with only small changes and the large steps.

Our model implies the existence of a mechanism that detects prediction error. Thus, it is consistent with the fact that the dopamine neurons signal prediction errors (Fiorillo, Tobler, & Schultz, 2003; Schultz et al., 1997; Schultz & Dickinson, 2000), a finding that has inspired temporal difference models (Rangel, Camerer, & Montague, 2008; Schultz, 2006). However, in our model this signal does not enter into a delta-rule update; rather it initiates a Bayesian model selection process that decides what has caused the error. The update that follows depends on how this decision plays out. Thus, our model offers an alternative interpretation for the function of the dopaminergic error signal.

Our model implies the existence of a second mechanism that must be, on average, more active when volatility is high (i.e., when changes occur frequently). Problems with a current estimate occur more often when there has been a change in the stochastic parameter than when there has not been. A change is detected only by the second stage computation. When it is detected, the new estimate is generally based on a small sample of recent observations. This means that the new estimate is often in substantial error, and so a problem with it is soon detected. There is neuroscientific evidence for a mechanism that becomes more active when volatility is high (Cohen et al., 2007; Yu & Dayan, 2005). Our model offers an alternative interpretation of the function of that mechanism: It mediates Bayesian model selection, rather than modulating a learning rate.

Our model estimates volatility, because it estimates  $p_c$ , the probability of a change in the value of the Bernoulli probability. This estimate affects its sensitivity to changes, because the estimate of  $p_c$  determines the prior odds of a change in a sequence of

any given length. When  $p_c$  is high, the prior odds of a change in any given sequence are higher; therefore, a Bayesian decision-making algorithm is more likely to decide that the sequence contains a change. The estimate of  $p_c$  is updated whenever the first stage detects a problem, before the second stage decides whether there has or has not been a change. The estimate of  $p_c$  exerts its effect on the sensitivity to change not by modulating a learning rate but rather by biasing the decision as to whether there has or has not been a change.

### Relevance to General Issues in Perception

Our model brings into prominence several general issues in perception. One of these is the effect of subsequent experience on the representation of previous experience. The trial-by-trial estimation experiment that inspired the model shows that future experience will alter the representation of past experience. Our model explains why this should be so: Information from observations not yet made is relevant to the question of whether another change point is justified. Therefore, the only way to build an accurate change-point description of the experienced sequence is to revise the description retrospectively as more information comes in. Our change-point model illustrates the kind of computational process that may underlie these retrospective alterations.

It is quite generally the case that the information in observations not yet made is relevant in deciding on the best description of what has already happened. Thus, generally speaking, if the brain is attempting to construct a compact description of its experience, on the implicit assumption that the information preserved in this description may prove useful in the future, revision of a previous representation in the light of later experience will frequently be observed.

When one revises a description that is constructed as experience unfolds, the issue of descriptive complexity cannot be avoided. Unless minimizing complexity (aka maximizing compactness or simplicity) is given its due, there is no reason not to represent  $p_g$  as having been 1 on every trial on which a green circle was observed and 0 on every trial on which a red circle was observed. In fact, this would be the correct description of the sequences in our experiment, because random number generation by a computer is a deterministic process. This very fact serves to emphasize the point that descriptive adequacy is more important than truth. Descriptive adequacy can only be achieved by applying Einstein's razor in adjudicating between adequacy and simplicity. A description should as simple as possible but not simpler. The comparison of marginal model likelihoods—Bayesian model selection—is a solution to this very general problem, as was first clearly realized by Jeffreys (1931).

We further suggest that representing the history of behaviorally important environmental parameters by perceived change points is a principle of broad application for information-theoretic reasons. Habituation to the steady state is observed everywhere in sensory signaling. It is seen even in the interoceptors that monitor homeostatic variables that change rarely. In his reviews of the literature on interoception, Dworkin (1993, 2007) called attention to a seeming paradox: Regulatory responses, both behavioral and physiological, appear to be based on the state of the variables being regulated, but the interoceptors sensitive to these variables signal only changes in state, not the states themselves. How, Dworkin

asked, does the brain know the steady state when it gets no steady signal indicative of it? The answer to this would appear obvious to a communications engineer: The information is in the changes; it is a waste of signal energy and channel capacity to steadily signal a steady state. Provided that the receiver has an updatable memory, signals specifying the magnitude and direction of changes are all it needs to maintain in memory a representation of the current state. Our model for how the brain represents an experienced Bernoulli sequence is an example of what we suggest is a more general principle governing the brain's representation of behaviorally important parameters of the experienced world: What the brain notes and remembers are the changes; from the coordinates of the change points, the current or any earlier state may be computed whenever it is needed. Moreover, summary statistics on change itself—its frequency and the distribution of change magnitudes—may be as readily computed.

A signature of representation by change points is the discontinuity in perception that our subjects show in their trial-to-trial perception of  $p_g$ . This discontinuity manifests itself behaviorally in the fact that subjects' perception of  $p_g$  sometimes changes by large amounts from one trial to the next. Large, maximally abrupt changes are not unique to this paradigm, nor to the behavior of human subjects. Similarly abrupt changes are observed in the free-operant matching behavior of rats and mice (Gallistel et al., 2007; Gallistel, Mark, King, & Latham, 2001). In the free-operant matching paradigm, the stochastic parameters on which behavior depends are the parameters of two concurrent Poisson processes. An abrupt (one-trial) adjustment to a change in Bernoulli probability is also seen in the distribution of switch latencies in a timing paradigm in which the target latency for a switch from one option to another depends on the relative frequency of the options (Balci et al., 2009; Kheifets & Gallistel, 2012). We take these large maximally abrupt changes to reflect the perception of a change in the stochastic parameter(s) and the consequent reestimation of their values, using only the data after the point in the past at which the change is estimated to have occurred.

Maximally large and abrupt behavioral changes cannot occur when the process that estimates the stochastic parameters governing behavior updates its estimates trial by trial. In all such models, the estimate on any given trial averages over some number of preceding outcomes. When there is a step change, the range of preceding outcomes entering into the estimate necessarily straddles the step. The straddle limits the maximum change that can occur from one trial to the next. By contrast, there is no limit on step height in the kind of model we propose, because the parameter estimates before and after an estimated change point are often based on nonoverlapping data sets.

Adjudicating between competing representations is inherently computationally demanding. For that reason, we suggest that the principle "If it ain't broke, don't fix it" has broad application. A less casual statement of this principle is that if there is no reason to doubt the current representation, do not trouble to decide among the alternatives to it. Deciding on an alternative is only undertaken when the current representation appears incompatible with recent observations, because it is easier to detect a fault in the current representation than to decide on an appropriate alternative to it.

Our model also illustrates the representation of different kinds of uncertainties, which has been the focus of recent theoretical work (Cohen et al., 2007; Yu & Dayan 2005). This recent work

distinguishes between “expected” uncertainties and “unexpected” uncertainties. To illustrate this distinction, Cohen et al. (2007) referred to the conditions of our experiment; namely, observing a nonstationary Bernoulli sequence. What they called the expected uncertainty is the uncertainty about whether the next circle will be green or red, given the current estimate of  $p_g$ . The closer  $p_g$  is to .5, the greater this uncertainty, because the entropy of a Bernoulli distribution is maximal when  $p = .5$ . Cohen et al. called this the expected uncertainty because they assume that subjects understand the extent to which the color of the next circle is unpredictable. What they called the unexpected uncertainty is the uncertainty that arises from nonstationarity, the possibility that  $p_g$  may change. They called this the unexpected uncertainty, because they implicitly assumed that subjects do not expect nonstationarity in stochastic parameters. Whether their assumption is true or not is debatable—at least some subjects may believe that nothing is certain but death and taxes—but it does not really matter. In Robinson’s experiment and ours, subjects expected that  $p_g$  might change, because we told them it might. The important point is that uncertainty about  $p_c$  must be distinguished from uncertainty about  $p_g$ . In our model, this distinction is realized in the distinction between the distribution on  $p_g$  and the distribution on  $p_c$ . The entropy of the first is the subject’s “expected” uncertainty; the entropy of the second is the subject’s “unexpected” uncertainty.

To our knowledge, our model is the first model of change detection in the psychological literature to take account of the growth of the prior odds of a change as the length of a sequence increases. The assumption that the computational process mediating the perception of probability takes account of this analytic truth may shed light on the gambler’s fallacy (cf. Brown & Steyvers, 2007). The gambler’s fallacy is that the longer a streak of bad luck grows, the more likely the streak is to come to an end. We suggest that this fallacy may arise from a strong a priori conviction that stochastic processes are stepwise nonstationary, together with a failure to appreciate the subtle difference between “more likely to come to an end” and “more likely to have come to an end.”

If a stochastic parameter is nonstationary, if it ever changes, then it is an analytic truth that the longer the sequence one has observed the more likely it is to contain a change. However, that change may not yet be manifest. That is, if one has had a long streak of bad cards and if the underlying stochastic process is nonstationary, the odds that the streak has already come to an end do increase as the streak gets longer. Thus, the streak may already have ended but that may not yet have become apparent. Part of the fascination of the problem of detecting a change in a hidden Markov parameter is that a change can in principle only be detected some number of observations after it has occurred. This makes the distinction between “more likely to *have* come to an end” and “more likely to *come* to an end” subtle. The probability that the change has already occurred but has not yet become apparent does increase with sequence length. What does not change—assuming that  $p_c$  is itself stationary—is the probability that the change will occur before the next observation.

Because our model assumes that the computation by which the brain estimates current probabilities and detects changes in probability has incorporated in it the analytic fact that the prior odds of a change increase with sequence length, it provides a psychological foundation for an explanation of the gambler’s fallacy. In typical examples of the gambler’s fallacy, the underlying stochas-

tic process is stationary, but those who fall prey to the fallacy may not represent it as such. They may deeply believe in changes in luck (or in transient changes in athletic prowess). For the general run of stochastic processes, this belief may be well founded; that is, it may be that most behaviorally important stochastic processes are nonstationary. For an alternative approach to modeling the origins of the gambler’s fallacy, see Brown and Steyvers (2009).

Finally, Robinson found, and we find, that under our experimental conditions neither the accuracy nor the precision of subjects’ estimates of the experienced probability varies as a function of the observed probability, over the full range of probabilities. Our finding accords with a number of reports going back more than 60 years (Fox & Hadar, 2006; Peterson & Beach, 1967; Ungemach, Chater, & Stewart, 2009), but it runs contrary to some other reports (for reviews, see Gonzalez & Wu, 1999; Luce, 2000). Atneave (1953), for example, had people estimate the frequencies of English letters. His subjects overestimated the frequency of rare letters and underestimated the frequency of common letters, quite dramatically.

The assumption that subjective probability is an ogival or inverse ogival function of objective probability has played a prominent role in the judgment and decisions literature ever since the seminal work of Kahneman and Tversky (1979). It has been suggested that the distortion is only seen when people are working from described (or single-event) probabilities, not experienced probabilities (Gigerenzer, 1994; Hertwig, Barron, Weber, & Erev, 2004), but there are reasons to doubt that this can be a general explanation (Hadar & Fox, 2009). For one thing, Atneave’s subjects were presumably judging letter frequencies from extensive experience.

In considering how to reconcile wildly discrepant findings on the form of the function relating subjective probability to objective probability, it is important to bear in mind that different participants in the same experiment may show functions with different forms (Gonzalez & Wu, 1999; Luce, 2000). Thus, functions based on cross-subject averages may have little meaning (Estes, 1956, 2002; Estes & Maddox, 2005). Even more important is the finding that the same subjects produce different functions in different tasks (Brooke & MacRae, 1977; Wu, Delgado, & Maloney, 2009). These findings would seem to require that we distinguish between the perception of the probability and task-specific transformations of the percept. Zhang and Maloney (2012) showed that across a wide variety of tasks, performance is well described by the assumption that it is based on a linear function of the log of the odds, with task-specific slope and intercept parameters. This suggestion brings coherence to an otherwise puzzling array of seemingly contradictory findings. Their suggestion would be consistent with the assumption that there is a veridical mapping from observed probability to perceived probability and that it is this perceived probability that is the argument of the task-specific linear-logit transformation that subjects apply in a variety of tasks.

In considering the relation between our finding that the perception of sequentially experienced nonstationary Bernoulli probability is veridical over the full range and the many demonstrations of radically suboptimal reasoning about and utilization of probabilities, it is important to distinguish between the ideal observer and the ideal agent. An observer is ideal insofar as it constructs a representation of some aspect of the world that is as veridical as is in principle possible, given the information that the observer has

been given. An ideal agent, given a representation with a specified degree of uncertainty about the state of the world, applies a strategy based on that representation that optimizes some function, such as the percent correct. The agent is ideal insofar as the strategy it deploys is the best strategy given that representation and that goal.

There are cases in which subjects appear to be approximately ideal observers but are far from ideal agents. One striking case is the discrete-trials probability matching paradigms in which subjects attempt to predict the next outcome of a Bernoulli process. Subjects tend to match the frequency with which they predict an outcome to the frequency with which it occurs (Maddox, 2004; Myers, 1976; Vulkan, 2000). Insofar as their prediction probability matches the probability estimate of the ideal observer, they are ideal observers. That is, they appear to correctly estimate the probability of a given outcome. However, on the assumption that they are trying to maximize the number of correct predictions, subjects who probability match are not ideal agents. The ideal agent, when assigned the task of maximizing correct predictions, would predict the more probable outcome on every trial. If subjects were ideal agents in a probability matching protocol, one would not be able to judge from their behavior the extent to which they were ideal observers. This somewhat paradoxical fact serves to emphasize the importance of making this distinction. The distinction is relevant to the question of whether and under what conditions psychological processes are optimal (Bowers & Davis, 2012).

## References

- Adams, R. P., & MacKay, D. J. C. (2006). *Bayesian online changepoint detection*. Retrieved from <http://arxiv.org/abs/0710.3742>.
- Atneave, F. (1953). Psychological probability as a function of experienced frequency. *Journal of Experimental Psychology*, *46*, 81–86. doi:10.1037/h0057955
- Balci, F., Freestone, D., & Gallistel, C. R. (2009). Risk assessment in man and mouse. *Proceedings of the National Academy of Science, USA*, *106*, 2459–2463. doi:10.1073/pnas.0812709106
- Balci, F., Freestone, D., Simen, P., deSouza, L., Cohen, J. D., & Holmes, P. (2011). Optimal temporal risk assessment. *Frontiers in Integrative Neuroscience*, *5*, Article 55. doi:10.3389/fnint.2011.00056
- Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*, 1214–1221. doi:10.1038/nn1954
- Bowers, J. S., & Davis, C. J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin*, *138*, 389–414. doi:10.1037/a0026450
- Brooke, J. B., & MacRae, A. W. (1977). Error patterns in the judgment and production of numerical proportions. *Perception & Psychophysics*, *21*, 336–340. doi:10.3758/BF03199483
- Brown, S. D., & Steyvers, M. (2009). Detecting and predicting changes. *Cognitive Psychology*, *58*, 49–67. doi:10.1016/j.cogpsych.2008.09.002
- Carandini, M. (2012). From circuits to behavior: A bridge too far? *Nature Neuroscience*, *15*, 507–509. doi:10.1038/nn.3043
- Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*, 933–942. doi:10.1098/rstb.2007.2098
- Cordes, S., Gallistel, C. R., Gelman, R., & Latham, P. (2007). Nonverbal arithmetic in humans: Light from noise. *Perception & Psychophysics*, *69*, 1185–1203. doi:10.3758/BF03193955
- Cordes, S., Gelman, R., Gallistel, C. R., & Whalen, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin & Review*, *8*, 698–707. doi:10.3758/BF03196206
- Corrado, G. S., Sugrue, L. P., Seung, H. S., & Newsome, W. T. (2005). Linear-nonlinear-Poisson models of primate choice dynamics. *Journal of the Experimental Analysis of Behavior*, *84*, 581–617. doi:10.1901/jeab.2005.23-05
- Courville, A. C., Daw, N. D., Gordon, G. J., & Touretzky, D. S. (2004). Model uncertainty in classical conditioning. In S. Thrun, L. Saul, & B. Schölkopf (Eds.), *Advances in neural information processing systems 16* (pp. 977–984). Cambridge, MA: MIT Press.
- Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences*, *10*, 294–300. doi:10.1016/j.tics.2006.05.004
- Dayan, P., & Daw, N. D. (2008). Decision theory, reinforcement learning, and the brain. *Cognitive, Affective & Behavioral Neuroscience*, *8*, 429–453. doi:10.3758/CABN.8.4.429
- Dayan, P., & Niv, Y. (2008). Reinforcement learning: The good, the bad and the ugly. *Current Opinion in Neurobiology*, *18*, 185–196. doi:10.1016/j.conb.2008.08.003
- Dworkin, B. R. (1993). *Learning and physiological regulation*. Chicago, IL: University of Chicago Press.
- Dworkin, B. R. (2007). Interoception. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (3rd ed., pp. 482–506). New York, NY: Cambridge University Press. doi:10.1017/CBO9780511546396.021
- Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin*, *53*, 134–140. doi:10.1037/h0045156
- Estes, W. K. (1957). Theory of learning with constant, variable, or contingent probabilities of reinforcement. *Psychometrika*, *22*, 113–132. doi:10.1007/BF02280048
- Estes, W. K. (1984). Global and local control of choice behavior by cyclically varying outcome probabilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 258–270. doi:10.1037/0278-7393.10.2.258
- Estes, W. K. (2002). Traps in the route to models of memory and decision. *Psychonomic Bulletin & Review*, *9*, 3–25. doi:10.3758/BF03196254
- Estes, W. K., & Maddox, W. T. (2005). Risks of drawing inferences about cognitive processes from model fits to individual versus average performance. *Psychonomic Bulletin & Review*, *12*, 403–408. doi:10.3758/BF03193784
- Fearnhead, P., & Liu, Z. (2007). On-line inference for multiple changepoint problems. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, *69*, 589–605. doi:10.1111/j.1467-9868.2007.00601.x
- Feldman, J. (2009). Bayes and the simplicity principle in perception. *Psychological Review*, *116*, 875–887.
- Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003, March 21). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, *299*, 1898–1902. doi:10.1126/science.1077349
- Fox, C. R., & Hadar, L. (2006). “Decisions from experience” = sampling error + prospect theory: Reconsidering Hertwig, Barron, Weber & Erev (2004). *Judgment and Decision Making*, *1*, 159–161.
- Gallistel, C. R. (2009). The importance of proving the null. *Psychological Review*, *116*, 439–453. doi:10.1037/a0015251
- Gallistel, C. R., & Gelman, R. (2005). Mathematical cognition. In K. Holyoak & R. Morrison (Eds.), *Cambridge handbook of thinking and reasoning* (pp. 559–588). New York, NY: Cambridge University Press.
- Gallistel, C. R., King, A. P., Gottlieb, D., Balci, F., Papachristos, E. B., Szalecki, M., & Carbone, K. S. (2007). Is matching innate? *Journal of the Experimental Analysis of Behavior*, *87*, 161–199. doi:10.1901/jeab.2007.92-05

- Gallistel, C. R., Mark, T. A., King, A. P., & Latham, P. E. (2001). The rat approximates an ideal detector of changes in rates of reward: Implications for the law of effect. *Journal of Experimental Psychology: Animal Behavior Processes*, *27*, 354–372. doi:10.1037/0097-7403.27.4.354
- Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is important for psychology and vice versa. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 129–161). New York, NY: Wiley.
- Glimcher, P. W. (2003a). *Decisions, uncertainty and the brain: The science of neuroeconomics*. Cambridge, MA: MIT Press.
- Glimcher, P. W. (2003b). The neurobiology of visual-saccadic decision making. *Annual Review of Neuroscience*, *26*, 133–179. doi:10.1146/annurev.neuro.26.010302.081134
- Glimcher, P. W. (2009). Choice: Towards a standard back-pocket model. In P. W. Glimcher, C. F. Camerer, E. Fehr, & R. A. Poldrack (Eds.), *Neuroeconomics: Decision making and the brain* (pp. 503–521). New York, NY: Academic Press.
- Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, *38*, 129–166. doi:10.1006/cogp.1998.0710
- Hadar, L., & Fox, C. R. (2009). Information asymmetry in decision from description versus decision from experience. *Judgment and Decision Making*, *4*, 317–325.
- Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choices. *Psychological Science*, *15*, 534–539. doi:10.1111/j.0956-7976.2004.00715.x
- Jeffreys, H. (1931). *Scientific inference*. New York, NY: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291. doi:10.2307/1914185
- Kakade, S., & Dayan, P. (2002). Acquisition and extinction in autoshaping. *Psychological Review*, *109*, 533–544. doi:10.1037/0033-295X.109.3.533
- Kheifets, A., & Gallistel, C. R. (2012). Mice take calculated risks. *Proceedings of the National Academy of Sciences, USA*, *109*, 8776–8779. doi:10.1073/pnas.1205131109
- Krugel, L. K., Biele, G., Mohr, P. N. C., Li, S.-C., & Heekeren, H. R. (2009). Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proceedings of the National Academy of Sciences*, *106*, 17951–17956. doi:10.1073/pnas.0905191106
- Luce, R. D. (2000). *Utility of gains and losses: Measurement-theoretical and experimental approaches*. Mahwah, NJ: Erlbaum.
- MacKay, D. J. C. (2003). *Information theory, inference and learning algorithms*. New York, NY: Cambridge University Press.
- Maddox, W. T. (2004). Probability matching, accuracy maximization, and a test for the optimal classifier's independence assumption in perceptual categorization. *Perception & Psychophysics*, *66*, 104–118. doi:10.3758/BF03194865
- Marr, D. (1982). *Vision*. San Francisco, CA: Freeman.
- Mathys, C., Daunizeau, J., Friston, K. J., & Stephan, K. E. (2011). A Bayesian foundation of individual learning under uncertainty. *Frontiers in Human Neuroscience*, *5*, Article 39. doi:10.3389/fnhum.2011.00039
- Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience*, *16*, 1936–1947.
- Myers, J. (1976). Probability learning and sequence learning. In W. Estes (Ed.), *Handbook of learning and cognitive processes: Approaches to human learning and motivation* (pp. 171–205). Hillsdale, NJ: Erlbaum.
- Nassar, M. R., & Gold, J. I. (2013). A healthy fear of the unknown: Perspectives on the interpretation of parameter fits from computational models in neuroscience. *PLoS Computational Biology*, *9*(4): e1003015. doi:10.1371/journal.pcbi.1003015
- Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasley, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature Neuroscience*, *15*, 1040–1046. doi:10.1038/nn.3130
- Nassar, M. R., Wilson, R. C., Heasley, B., & Gold, J. I. (2010). An approximately Bayesian delta-rule model explains the dynamics of belief updating in changing environment. *Journal of Neuroscience*, *30*, 12366–12378. doi:10.1523/JNEUROSCI.0822-10.2010
- Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*, 139–154. doi:10.1016/j.jmp.2008.12.005
- Peterson, C. R., & Beach, L. R. (1967). Man as an intuitive statistician. *Psychological Bulletin*, *68*, 29–46. doi:10.1037/h0024722
- Polunchenko, A. S., & Tartakovsky, A. G. (2012). State-of-the-art in sequential change-point detection. *Methodology and Computing in Applied Probability*, *14*, 649–684. doi:10.1007/s11009-011-9256-5
- Preuschhoff, K., & Bossaerts, P. (2007). Adding prediction risk to the theory of reward learning. In B. W. Balleine, K. Doya, J. O'Doherty, & M. Sakagami (Eds.), *Annals of the New York Academy of Sciences: Vol. 1104. Reward and decision making in corticobasal ganglia networks* (pp. 135–146). doi:10.1196/annals.1390.005
- Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*, 545–556. doi:10.1038/nrn2357
- Rapoport, A., & Wallsten, T. S. (1972). Individual decision behavior. *Annual Review of Psychology*, *23*, 131–176. doi:10.1146/annurev.ps.23.020172.001023
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64–99). New York, NY: Appleton-Century-Crofts.
- Robinson, G. H. (1964). Continuous estimation of time-varying probability. *Ergonomics*, *7*, 7–21. doi:10.1080/00140136408930721
- Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology*, *57*, 87–115. doi:10.1146/annurev.psych.56.091103.070229
- Schultz, W., Dayan, P., & Montague, P. R. (1997, March 14). A neural substrate of prediction and reward. *Science*, *275*, 1593–1599. doi:10.1126/science.275.5306.1593
- Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience*, *23*, 473–500. doi:10.1146/annurev.neuro.23.1.473
- Seymour, B., O'Doherty, J. P., Dayan, P., Koltzenburg, M., Jones, A. K., Dolan, R. J., . . . Frackowiak, R. S. (2004, June 10). Temporal difference models describe higher-order learning in humans. *Nature*, *429*, 664–667. doi:10.1038/nature02581
- Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual Review of Psychology*, *28*, 1–39. doi:10.1146/annurev.ps.28.020177.000245
- Steyvers, M., & Brown, S. (2006). Prediction and change detection. *Advances in Neural Information Processing Systems*, *18*, 1281–1288.
- Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2005). Choosing the greater of two goods: Neural currencies for valuation and decision making. *Nature Reviews Neuroscience*, *6*, 363–375. doi:10.1038/nrn1666
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning*. Cambridge, MA: MIT Press.
- Turner, R., Saatici, Y., & Rasmussen, C. E. (2009). *Adaptive sequential Bayesian change point detection*. Retrieved from <http://mlg.eng.cam.ac.uk/rturner/BOCPD.pdf>
- Ungemach, C., Chater, N., & Stewart, N. (2009). Are probabilities overweighted or underweighted when rare outcomes are experienced (rarely)? *Psychological Science*, *20*, 473–479. doi:10.1111/j.1467-9280.2009.02319.x

- Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys*, *14*, 101–118. doi:10.1111/1467-6419.00106
- Wallsten, T. S., & Budescu, D. V. (1983). Encoding subjective probabilities: A psychological and psychometric review. *Management Science*, *29*, 151–173. doi:10.1287/mnsc.29.2.151
- Wilson, R. C., Nassar, M. R., & Gold, J. I. (2010). Bayesian online learning of the hazard rate in change-point problems. *Neural Computation*, *22*, 2452–2476. doi:10.1162/NECO\_a\_00007
- Wu, S.-W., Delgado, M. R., & Maloney, L. T. (2009). Economic decision-making compared with an equivalent motor task. *Proceedings of the National Academy of Sciences, USA*, *106*, 6088–6093. doi:10.1073/pnas.0900102106
- Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation and attention. *Neuron*, *46*, 681–692. doi:10.1016/j.neuron.2005.04.026
- Zhang, H., & Maloney, L. T. (2012). Ubiquitous log odds: A common representation for probability and frequency distortion in perception, action, and cognition. *Frontiers in Neuroscience*, *6*, Article 1.

## Appendix A

### Experimental Method

Ten subjects were recruited from a diverse population of undergraduate seniors, graduate students, and postgraduates, between 20 and 30 years of age, from a variety of disciplines such as computer science, psychology, and business. They included native speakers of English, Russian, Chinese, and Bulgarian. Female and male subjects were equal in number. They were paid \$10 per session. They sat in front of a 21-in. Macintosh monitor. A cartoon of the screen is shown in Figure 3 of the main text. The following instructions appeared on the screen in the practice version of the experiment:

On the following slide you will see a box labeled “RINGS.” This box contains an unknown number of red and green rings.

Your task is to estimate the percentage of GREEN rings in this box based on a small sample of rings shown to you, one at a time.

From time to time, without warning, the box of rings will be replaced with a NEW box. You will not be told that a change has occurred.

If you believe the box has been changed you should indicate that a change has occurred and then continue estimating the percentage of GREEN rings in the new box.

#### PROCEDURE

1. Begin with a guess. Adjust the position of the slider to indicate your estimate. Click the button labeled “NEXT” for the next trial.
2. Thereafter, on each trial you will be shown a single ring chosen randomly from the box of RINGS.
3. You may then change your estimate of the percentage of GREEN rings in the box or you may leave your estimate unchanged. Again, click “NEXT.”
4. Once you adjust the slider to your liking and click “NEXT, the ring will be RETURNED to its box.

5. This process will be repeated on each trial.
6. If you think the box has been changed click the button labeled “I think the box has changed.”
7. This will commence the next trial and you will continue as before.

**FOR YOUR CONVENIENCE:** Your current estimate of the percentage of GREEN rings will be reflected in the distribution of RED and GREEN circles shown on the right-hand side of the screen. This will enable you to “see” your chosen percentage of GREEN rings.

**NOTE:** Since the ring is always returned to its box, a particular ring may be drawn from the box on more than one trial.

Click “BEGIN” to start your practice session.

When we discovered from debriefing the first five subjects, and from running ourselves as subjects, that subjects generally had second thoughts, we added the following instructions, which were seen by the last 5 subjects:

8. If after a few trials you decide to take back your assertion of a change, click on the button labeled “I take that back!”
9. This will commence the next trial and you may continue as before.

The rectangle at upper right of the screen contained 1,000 circles, some red, some green. The proportion of green circles varied linearly from 0 to 1 in accord with the position to which the subject set the slider. This visual feedback may initially have been helpful to subjects, but many subjects indicated at the conclusion of the experiment that they paid little attention to it.

(Appendices continue)

## Appendix B

### Determining Whether the Current Estimate of $p_g$ (the Slider Setting) Needs to Be Changed

The first decision—and after most trials, the only decision—is whether the number of green circles observed since the last change point, denoted  $n_g$ , is consistent with the current estimate of the hidden probability, denoted  $\hat{p}_g$ . Here, consistent with means “not too far out on the tails of the probability distribution on  $n_g$  green circles.” Thus, we need to compute that probability distribution, which is binomial,

$$p(n_g | \hat{p}_g, n) = \hat{p}_g^{n_g} (1 - \hat{p}_g)^{n - n_g} \frac{n!}{n_g! (n - n_g)!}, \quad (\text{B1})$$

where  $n$  is the total number of circles observed since the last change point.

We treat  $n_g/n$  as a continuous variable and write down an approximate expression for it. Our starting point is to use Sterling’s formula to approximate the factorials, yielding, after a small amount of algebra,

$$P(n_g | \hat{p}_g, n) \propto \exp(-nD_{\text{KL}}(p_o \| \hat{p}_g)), \quad (\text{B2})$$

where  $D_{\text{KL}}(p_o \| \hat{p}_g)$  is the Kullback–Leibler divergence defined in Equation 2 of the main text, and  $p_o = n_g/n$  is the fraction of green circles observed since the last change point. We have replaced the discrete variable,  $n_g$ , with the continuous one on the left-hand side of Equation B1.

Equation B2 is still nontrivial, but if we make the approximation that  $n$  is reasonably large, the right-hand side of Equation 5 is nonnegligible only when  $p_o$  is close to  $\hat{p}_g$ , and we can Taylor expand  $D_{\text{KL}}(p_o \| \hat{p}_g)$  around  $p_o = \hat{p}_g$ . When we do that,  $p(n_g | \hat{p}_g, n)$  becomes Gaussian in the difference  $p_o - \hat{p}_g$ , and, therefore,

$2nD_{\text{KL}}(p_o \| \hat{p}_g)$  becomes chi-squared with one degree of freedom; therefore, because the chi square distribution is the  $\Gamma(v/2, 2)$  distribution with  $v = \text{degrees of freedom}$ ,  $nD_{\text{KL}}$  is distributed  $\Gamma(0.5, 1)$ .

This analysis can be used to determine whether a set of observations is consistent with an estimate of the probability of a green circle. It can also be used to estimate how long we would have to wait to determine that a particular estimate is wrong: We simply replace  $p_o$  with the true probability of a green circle. In the main text, we make that replacement, except that we use the posterior mean of a Bayes optimal probability tracking algorithm in place of the true probability.

Finally, we show how to explicitly convert Kullback–Leibler divergences to conventional  $p$  values. Using the fact that  $\Gamma(0.5, 1)$  is closely related to the Gaussian, we have, after straightforward algebra,

$$P(2nD_{\text{KL}} > z) = 2 \int_{z^{1/2}}^{\infty} \frac{e^{-x^2/2}}{(2\pi)^{1/2}} = 2[1 - \Phi(z^{1/2})],$$

where  $\Phi(z)$  is the cumulative normal function. Alternatively we may write

$$P(nD_{\text{KL}} > z) = 2[1 - \Phi((2z)^{1/2})].$$

The right-hand side is just alpha in a conventional test. For two common alphas used in the main text, 0.05 and 0.5 (the latter corresponding to even odds), the above expression tells us that  $z = 1.92$  and 0.22, respectively.

*(Appendices continue)*

## Appendix C

## Bayes Optimal Probability Tracking

The behavioral experiment described in the main text can be cast as a hidden Markov model: There is an unobserved true probability of a green circle,  $p_g$ , which may change on any trial. We observe green and red circles, and we want to infer the probability distribution over  $p_g$ . In other words, we want to infer  $P_t(p_g | x_1, x_2, \dots, x_t)$  where  $t$  labels trial and  $x_t$  tells us the color of the circle on trial  $t$ . For convenience, we use a convention in which  $x_t = 1$  if there is a green circle on trial  $t$  and  $x_t = 0$  if there is a red circle.

The probability that  $p_g$  changes on any trial is  $p_c$ , which is unknown. We thus have to average over it, and so we have

$$P(p_g | x_1, x_2, \dots, x_t) = \int dp_c P_t(p_c, p_g | x_1, x_2, \dots, x_t), \quad (C1)$$

where  $P_t(p_c, p_g | x_1, x_2, \dots, x_t)$  is the probability distribution over both  $p_c$  and  $p_g$  given the observations up to time  $t$ . This expression suggests that we should compute  $P_t(p_c, p_g | x_1, x_2, \dots, x_t)$  and, once we know that, integrate over  $p_c$  to find  $P_t(p_g | x_1, x_2, \dots, x_t)$ .

The formalism for doing this is standard. To simplify our equations we adopt the notation

$$x_{1:t} \equiv x_1, x_2, \dots, x_t.$$

Then, using the Bayes theorem, we have

$$P_t(p_c, p_g | x_{1:t}) \propto P(x_t | p_c, p_g, x_{1:t-1}) P_t(p_c, p_g | x_{1:t-1}). \quad (C2)$$

The probability of observing a green circle depends only on the underlying probability,  $p_g$ , so the first term simplifies to  $P(x_t | p_g)$ . For the second, we use

$$P_t(p_c, p_g | x_{1:t-1}) = \int dp'_c dp'_g P_t(p_c, p_g | p'_c, p'_g) P_{t-1}(p'_c, p'_g | x_{1:t-1}), \quad (C3)$$

where  $P(p_c, p_g | p'_c, p'_g)$  is the probability distribution over  $p_c$  and  $p_g$  on trial  $t$ , given that we know  $p'_c$  and  $p'_g$  on trial  $t - 1$ , and we haven't observed any new data.

In our experiments,  $p_c$  doesn't change from trial to trial, so the  $p_c$ -dependent piece of the conditional probability is just a Dirac delta function,  $\delta(p_c - p'_c)$ . For the  $p_g$ -dependent piece, we use the fact that with probability  $1 - p_c$ ,  $p_g$  doesn't change, and with probability  $p_c$ , it is drawn from the prior. We thus have

$$P(p_c, p_g | p'_c, p'_g) = \delta(p_c - p'_c) [(1 - p_c) \delta(p_g - p'_g) + p_c P(p_g)],$$

where  $P(p_g)$  is the prior probability distribution over  $p_g$ . Inserting this into Equation C3, inserting the resulting expression into Equation C2, and using the fact that  $P(x_t | p_c, p_g, x_{1:t-1}) = P(x_t | p_g)$ , we have

$$P_t(p_c, p_g | x_{1:t}) \propto P(x_t | p_g) \left[ (1 - p_c) P_{t-1}(p_c, p_g | x_{1:t-1}) + p_c P(p_g) \right].$$

Finally, we note that  $P(x_t | p_g)$  is Bernoulli,

$$P(x_t | p_g) = p_g^{x_t} (1 - p_g)^{1-x_t}.$$

This leads to the expression

$$P_t(p_c, p_g | x_{1:t-1}) = \frac{p_g^{x_t} (1 - p_g)^{1-x_t} \left[ (1 - p_c) P_{t-1}(p_c, p_g | x_{1:t-1}) + p_c P(p_g) P_{t-1}(p_c, p_g | x_{1:t-1}) \right]}{\int dp_c dp_g p_g^{x_t} (1 - p_g)^{1-x_t} \left[ (1 - p_c) P_{t-1}(p_c, p_g | x_{1:t-1}) + p_c P(p_g) \right]}, \quad (C4)$$

where we have explicitly included the normalization.

To solve Equation C4 (which must be done numerically), we need to specify a prior distribution over  $p_c$  and  $p_g$ . For simplicity—and because we found it didn't make much difference—we assume the prior over  $p_c$  is a delta function centered on the true change probability, denoted  $p_{c0}$ ,

$$P(p_c) = \delta(p_c - p_{c0}).$$

For the prior over  $p_g$  we use, as in the main text, a Beta distribution,

$$P(p_g) = \frac{p_g^{\alpha-1} (1 - p_g)^{\beta-1}}{B(\alpha, \beta)},$$

where  $B(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha + \beta)$  is the Beta function. Because of the delta function prior on  $p_c$ , Equation C4 simplifies to

$$P_t(p_g | x_{1:t-1}) = \frac{p_g^{x_t} (1 - p_g)^{1-x_t} \left[ (1 - p_{c0}) P_{t-1}(p_g | x_{1:t-1}) + p_{c0} P(p_g) \right]}{\int dp_g p_g^{x_t} (1 - p_g)^{1-x_t} \left[ (1 - p_{c0}) P_{t-1}(p_g | x_{1:t-1}) + p_{c0} P(p_g) \right]}. \quad (C5)$$

Note also that we no longer need to do the integral in Equation C2; instead, we just set  $p_c$  to  $p_{c0}$ . Equation C5 is what we used in the main text to compute the posterior mean of the Bayes-optimal probability tracking algorithm. C-code implementing the computation is given in the supplemental material.

Received July 13, 2012

Revision received August 13, 2013

Accepted August 16, 2013 ■