TIME AND PROBABILITY IN THE MOUSE

By AARON KHEIFETS

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

Graduate Program in Psychology

written under the direction of

Charles Randy Gallistel, Ph.D.

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

January 2014

ABSTRACT OF THE DISSERTATION

Time and Probability in the Mouse

By AARON KHEIFETS

Dissertation Director:
C. Randy Gallistel

I describe a series of four experiments studying decision under uncertainty

in the mouse and argue that their results demand a more sophisticated,

information-processing, account of mouse behavior than current popular models.

The tasks are based on the switch procedure of Balci, Freestone and Gallistel (2009),

which employs interval-timing behavior to study decision under uncertainty. I show

that: 1. Subjects respond to changes in task-relevant probabilities abruptly. 2. The

number of trials before they detect a change in probability is predicted by the

Kullback-Leibler divergence between the Bernoulli distributions of the two

probabilities. 3. Subjects are capable of decreasing their timing variability. 4.

Subjects differentiate between added exogenous temporal variability and their

endogenous timing variability. I argue that these four results are best understood in

an information-processing framework and modeling mouse behavior with

algorithms that explicitly represent abstract quantities such as probability,

exogenous variability and objective time.

Table of Contents

Introduction

In this thesis I report four experiments on time-based switching decisions in mice. I argue that the results of these experiments are relevant to a much broader discussion about judgments and decisions in other animals, including humans. I argue that the four experiments reported here demonstrate that:

1. Subjects *detect* changes in the probabilities integral to their task; rather than gradually adapting their behavior to whatever the frequency of events has been recently.

2. The speed with which they detect such changes is best characterized in information-theoretic terms.

3. Subjects can make their timing behavior more precise when the task demands it.

4. Subjects differentiate between sources of variability that are endogenous and those that are exogenous.

I will argue that these results support a model of the decision processes of mice that is probabilistically sophisticated, essentially concerned with the processing of information. This runs contrary to popular views of animal decision (and non-conscious human decision) that suggest subjects merely imitate the performance of a mathematically sophisticated decision-maker via the use of simple heuristics. In this introduction, I try to place these experiments in the broad context in which I believe they are relevant.

One central aim of this thesis is to demonstrate the virtues of an information-processing account of animal decision-making. The dichotomy between the type of thinking that underlies animal behavior and the type of explicit verbal reasoning that humans are capable of is at least as old as Plato. He, and the vast majority of thinkers following him, cast careful, deliberate reasoning as the path to truth, happiness and The Good. By contrast then, the quick, automatic and instinctual system that motivates animal behavior and unenlightened human behavior is classically seen as an inferior process that we use when there is insufficient time or mental resources for proper reasoning. In the last 35 years, research has shown this is a fiction: a large portion of the time, human reasoning is deeply flawed. Recent research, including this thesis, is finding that instinctual, "gut," decisions are more sophisticated than previously thought. This thesis shows that even mice are adept at calculating nearly optimal behavior in scenarios that require computing subtle probabilistic quantities and serves as an example of the use of an information-processing framework for understanding choices evident in animal behavior.

To be clear: I will advocate a view of animal decision-making that ascribes to animals the ability to estimate and perform calculations with complex and abstract probabilities. I will argue that the experiments in this thesis show behavior that cannot be explained without ascribing to our subjects (mice) a representation of probability values, likelihoods of higher-order events and even separate representations of uncertainty inherent in their sensory apparatus and uncertainty due to the environment.

The cracks in human rationality

In the 1970s Kahneman and Tversky published a series of studies that showed systematic errors in human subjects' responses to simple questions that dealt with uncertainty. For example, Tversky and Kahneman (1983) found that when subjects were told about a hypothetical person, Linda, who had traits that are typically associated with feminism, subjects rated the proposition "Linda is a bank teller and is active in the feminist movement" as more likely than "Linda is a bank teller." Obviously, this is a departure from probability theory because the conjunction of any two events $A \wedge B$ cannot be more likely than the probability of either event A or B taken individually. They found this effect to be extremely robust, even in subjects with training in statistics and even when the question was phrased to draw subjects' attention to the fact that in every circumstance in which Linda is a bank teller and a feminist, she is also a bank teller (but not the other way around). The most effective way to lower the proportion of subjects committing the conjunction fallacy was to change from asking about which events were more probable to asking subjects which proposition they should bet $10 on so as to maximize the probability of winning money. Even then, over half of participants still made the error.

Experts also made the conjunction fallacy, even in their domain of expertise. Tversky and Kahneman gave doctors a problem formally equivalent to the Linda problem above but about patients with a disease being more or less likely to experience each of a set of symptoms from a list. Once again, they found that their subjects evaluated the probability of the patient reporting symptoms $A \wedge B$ as more

likely than *B* when $A \wedge B$ was more typical of the disease. Thus even experts, with training in probability and statistics, engaging in content within their realm of expertise, making a judgment highly typical for their line of work, committed the conjunction fallacy.

The conjunction fallacy is just one example of incorrect reasoning Tversky and Kahneman (1974) found. Subjects also failed to take into account base-rates at which events occur, were insensitive to sample size and reliability considerations and were even swayed by the presentation of completely irrelevant numbers that were clearly identified as irrelevant. They found many more departures from correct reasoning in a wide variety of contexts, which spawned an entire field called Judgment and Decision Making full of examples of how human reasoning can yield illogical results.

These results do not necessarily mean that humans are deeply irrational, as Kahneman and Tversky point out (Tversky & Kahneman, 1986). They refer to these errors as "cognitive illusions" analogous to optical illusions. Though there are circumstances in which our visual system consistently perceives something other than what is actually presented, these situations do not mean that our visual system is deeply flawed. Despite knowing about optical illusions, we rightfully trust our visual system because the circumstances that produce illusions are rare and fragile. When we walk around in the world, the probability of encountering illusions in important circumstances and for an extended period of time is so low that they are negligible.

The study of cognitive illusions may give us insight into how we made decisions under uncertainty just as visual illusions give us insight into how the visual system works. The bugs of the visual system give us clues as to how it functions. If we can be aware of these quirks, we can minimize their potential for harm. Accordingly, Kahneman and Tverksy developed Prospect Theory, which was supposed to be the general theory that would account for the cognitive illusions they had discovered and how they can be thought of as the aberrant cases of a typically functional rational system.

Unfortunately, Prospect Theory (Kahneman & Tversky, 1979) and its successor, Cumulative Prospect Theory (Tversky & Kahneman, 1992), has not yielded a successful general account of decision-making under uncertainty. This is for two reasons. First, the number of exceptions to its application seems to grow in almost one-to-one correspondence with the number of studies of it. The ever-increasing list of quirks and exceptions to effects in this area means that we are left without a general framework in which to understand which effects take precedence over which others in a complex decision outside of the lab.

Second, the "cognitive illusions" it describes do not seem rare and fragile. On the contrary, as divergence from optimality in decision-making has become better documented, there has been an explosion of marketing programs exploiting the heuristics and biases that supposedly underpin our human reasoning. We should be worried indeed if our much-exalted human capacity for reason is simply a mélange of biases and heuristics that so often lead us away from good decisions.

The virtues of instinct

The picture of human reasoning painted above is bleak. At the same time, people seem perfectly capable of making good decisions with limited information hundreds of times a day: Which grocery store is likely to be less crowded right now? Can I make it across the street before that car? Will this fish make me sick? The flaws in human reasoning explored in Judgment and Decision Making are surprising to researchers exactly because of the fact that we typically do not have difficulty the real world despite having very incomplete information about it. If we walked around making grievous errors with dire consequences more often than not, our departures from normative reasoning would be unremarkable (and we might not even live long enough to reflect on that fact).

Other experiments attest to the robustness of human decision-making, showing that in a wide variety of tasks subjects are able to make judgments reasonably well under uncertainty, using very simple rules of thumb or *heuristics*. For example, when asked to guess which of two cities has the higher population, subjects often choose the city that is familiar to them, which is a simple and highly effective strategy (Gigerenzer & Goldstein, 1996). This is referred to as the recognition heuristic and is one of the seminal findings in this area. Similarly valid responses are also found in probabilistic judgments when the relevant probabilities are experienced (in a sequence of trials) by the subject rather than merely described to them (see Hertwig & Erev, 2009 for a review). This has led some researchers to extoll the deep wisdom of the heuristics, noting that in a number of cases heuristics can yield good solutions more reliably than more complex "normative" strategies

(Gigerenzer & Gaissmaier, 2011). As Gigerenzer and Goldstone argue (2009), this is essentially the problem of overfitting, illustrated in Figure 1: Naïvely fitting a complex model to data can cause the model to fit the noise in the data. The additional complexity in the model gives it the freedom to fit aspects of the data that are spurious. This makes it a better fit to the existing data but a worse predictor of future data generated by the same process. Thus, unnecessary model-complexity can make a model worse.
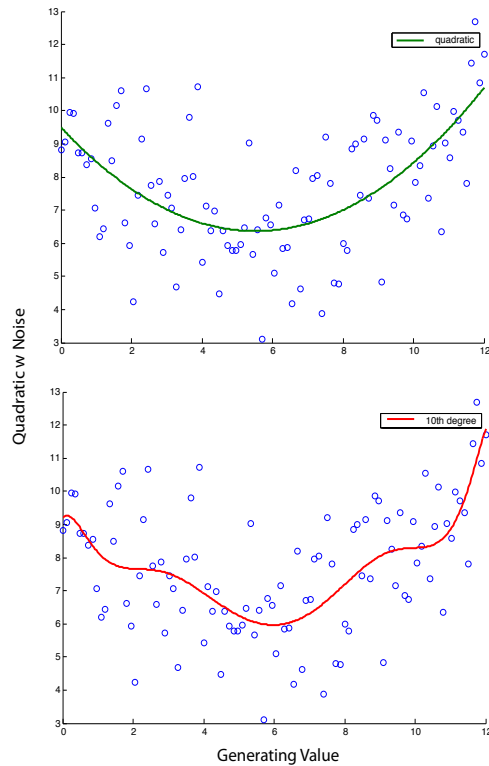
Figure 1. Here we see a set of toy data: generated from a parabola with Gaussian noise added. The upper panel displays the best fit to the data using a second degree polynomial and the lower panel displays the best fit to the same data using a tenth degree polynomial. Because the tenth degree polynomial has more degrees of freedom, it is a better fit to the data. Nevertheless, it will be a worse predictor of future data generated by the same process: since the generating function is approximately a parabola, the incoming data will tend to fall on that parabola, not this strange tenth degree polynomial. This phenomenon is called "overfitting" or "fitting the noise." Gigerenzer and colleagues claim that many "normative" solutions to experimental tasks are actually non-normative in a broader context, analogous to overfitting the data (Goldstein & Gigerenzer, 2009).

Given the traditional view that reason is the path to Truth and instinct is likely to lead us astray, we're left with a counterintuitive pattern of results. We can be terrifyingly irrational when asked to think about a situation described to us but surprisingly rational when we learn probabilities through experience. This can be seen as similar to the fact that we are able to accurately throw a ball to someone running over an uneven surface but most of us are unable to solve the set of

differential equations that would give us the normative solution for how to do this. Researchers studying skilled outfielders found that in attempting to catch a ball lobbed in the air, outfielders adjust their speed so that the angle of elevation of gaze to the ball is always at zero (McLeod & Dienes, 1996). Gigerenzer calls this the gaze heuristic because it does not require the subject to do complex differential equations or take into account difficult factors such as wind speed or the spin of the ball. He argues that we actually do catch balls using this heuristic (Gigerenzer, 2004) and that this is a prime example of how the simplicity of our instincts are actually their strength; making us essentially rational agents armed with a Swiss army knife of simple strategies selected by evolution for their robustness (the "adaptive toolbox").

There are, however, problems with this view. For one, some cornerstone heuristics have only been shown to be strategies that *could* give performance similar to human performance in quick, intuitive decisions. This is a world away from showing that humans actually *do* employ these heuristics. As pointed out by Hilbig (2010), insufficient attention has been paid to the fact that since heuristics have a high degree of validity and human subjects do well on these tasks, some degree of similarity in their outcomes is guaranteed.

For example, there is reason to doubt the use of the much-celebrated recognition heuristic. Hilbig (2010) points out several studies that show that subjects make the same choice as the recognition heuristic when it will give the correct answer but deviate from it more often when it will give the incorrect answer. Naturally, it would be impossible for subjects to do this if they were only

using the heuristic without other cues as to whether it will be correct on a particular question. Indeed, there is a sizeable body of evidence showing that subjects do not simply use the recognition heuristic, blind to other information. Perhaps unsurprisingly, subjects do not use this heuristic when they recognize cities because they are merely near where they happen to be or for other coincidental reasons such as knowing Nantucket because it is in a popular limerick (Oppenheimer, 2003).

A second problem with the "adaptive toolbox" view of rationality is that pointing out the wisdom of intuitions and simple strategies does not get us any closer to a unified framework in which to understand decisions. Much like the biases found with deliberate reasoning in the previous section, the heuristics that allegedly underpin our fast reasoning are merely listed by current theories rather than explained. For instance, there is no satisfying explanation of the strategy selection problem: when do subjects use which heuristic? (Glöckner, Betsch, & Schindler, 2010; Newell, 2005).

The current status of the field of Decision is similar to that ascribed by Joshua Tenenbaum (1999a) to the field of Categorization fifteen years ago: divided between the classical view, that categories are defined by sets of necessary and sufficient criteria (rules) and the statistical view, that categories are defined by exemplars or prototypes (similarity). He argued that a unified approach was necessary to understand how modules interact, why there are the number that there are, why they operate the way they do, and what they have in common. Indeed, the hotly contested question of whether categorization proceeds by exemplars or prototypes has been shown to be a false dichotomy; two points on a computationally important

spectrum (Briscoe & Feldman, 2011). It is my belief that such a general framework is necessary in the field of Decision to unify the disparate effects and theories (or lack thereof).

I think this can be done under a framework very similar to the one advocated by Tenenbaum and others (e.g. Chater, Tenenbaum, & Yuille, 2006; Griffiths, Chater, Kemp, Perfors, & Tenenbaum, 2010; Tenenbaum, 1999a), which I will call an Information Processing framework. I take this view in part because the tasks of decision under uncertainty and categorization are very similar (a categorization *is* a decision under uncertainty of "which category does this belong to?" and a decision under uncertainty is implicitly a categorization of options). It is also in part because this framework has already yielded useful results in the field, which will be discussed below.

Decision-Making from the Perspective of Information Processing

If the classic theories of Judgment and Decision Making were unable to explain diverse results and predict results in novel situations, it is perhaps time to take a step back and examine their assumptions. A key assumption, particularly with regard to models of intuitive decisions, has been that subjects are not performing complex probabilistic calculations because this would be too complicated. I argue that this assumption should be dropped. This is because of the usefulness of using an Information Processing framework when studying decisions, which sometimes entails ascribing a good deal of mathematical sophistication to our subjects. I will take a moment to discuss this assertion because the subjects in this thesis are all mice, which for some will make the claim of mathematical sophistication sound

strange. I will not, however, shy away from the assertion that subjects are performing sophisticated mathematical calculations despite the fact that they are mice.

I do not propose that mice secretly perform mathematical calculations with a pen and paper while researchers aren't looking. In the same way that a child does not need to know the laws of biology in order to grow into an adult, an animal does not need to consciously be aware of the mathematical formulae driving its decisions in order to make them. In addition, making decisions under uncertainty (or as it is more commonly known, "making decisions") is one of the fundamental activities of all animals. We have differing goals (don't get eaten, run down that prey, get to the cleaners before they close, etc.) but we're all plagued by the fact we don't know everything we'd like to. If we did, there would be no purpose to learning what cues predict favorable conditions: if dogs knew when food was coming and when it wasn't, any ringing of bells would be irrelevant to them. Of course, this is not the case: conditional learning can be seen in animals from humans to even single-celled paramecia (Armus, Montgomery, & Jellison, 2006). Making decisions with incomplete information is more universal in the animal kingdom than having eyes, lungs or a gender: it's what all animals do all day, every day.

It does not seem absurd, then, that animals would have a highly developed engine for evaluating decisions under uncertainty. This thesis combines perspectives from various disciplines—including, risk assessment, signal detection theory, and information processing—to show how recent, more conceptually sophisticated frameworks can yield more fruitful, simple, and predictive theories of

how intuitive judgments work. This is obviously a vital step in developing a unified theory of rationality and decision under uncertainty.

A brief word on complexity

This thesis will not dwell on questions of computational tractability for the computations I use to model subject behavior. This is not because I believe this question is unimportant but because such concerns presuppose that we understand the more fundamental processes in the brain such as the physical basis of memory which is itself not well understood at all (C.R. Gallistel & King, 2009). There is also no standard against which to compare the computational complexity of more complex calculations because there has been no definitive study of the complexity of selecting and implementing a set of heuristics working in conjunction. Given the number of heuristics proposed, the computational complexity of this task could be large indeed. Not only this, but existing models of decision are completely silent on how their inputs are calculated. They often take abstract quantities such as "familiarity" as primitive inputs, when some kind of statistical processing is clearly needed to get these inputs in the first place. Presuming further statistical processing (possibly of the same basic variety) may not add much computational complexity.

In addition, an Information Processing framework is useful regardless of how its functions are implemented. It provides us with an optimal standard against which we can compare our data. Since subjects approach optimality in a variety of tasks, there is a sense in which it is vacuously true that some implementation of my modeling functions (or an approximation thereof) must be performed by the subjects in some sense. On the other hand, if we ask whether decision under

uncertainty proceeds by algorithms such that each constituent part can be put into a one-to-one mapping with my model, with no shortcuts or approximations whatsoever, we can see this is certainly false. What I will argue in this thesis is that the functions being computed are sensitive to quantities that are abstract and subtle such as probability values, the probability of a probability value changing or the amount of variability that is due to the subject as opposed to the environment. Though it is difficult to make a sharp distinction between information processing and simple heuristics, the fact that subjects are sensitive to such abstract and complex quantities suggests that they are approximating normative information processing computations rather than greatly simplified heuristics.

The case for Information Processing

Most importantly, the experimental data have been highly encouraging for this approach. This section will very briefly survey some of the success of Bayesian models in describing human behavior from a variety of modalities and experimental methods. The goal here is not an exhaustive review but rather a few key examples that show the power of these models in diverse domains, especially non-verbal ones because those will be the focus of this thesis.

Bayesian models have been used to great effect in modeling human categorization (Tenenbaum, 1999b). As mentioned above, this is a topic deeply related to decision under uncertainty. For example, a hierarchical Bayesian model has been used to reconstruct phylogenetic trees from lists of traits each animal has and then to classify new animals (sets of traits like "has fur," "lays eggs," etc.) which made impressively similar classifications to subjects, including intuitive mistakes

such as classifying a whale as a fish (Kemp, Perfors, & Tenenbaum, 2007). Though

particularly evocative, this result is not atypical for Bayesian models of human

learning (for a review, see Tenenbaum, Kemp, Griffiths, & Goodman, 2011).

They have also been effective in modeling subjects' ability to make

reasonable predictions in everyday situations: subjects appear to have a reasonable

prior distribution on data they have encountered only in passing and appear to

combine independent sources of information with these priors and with one

another so as to make accurate predictions about commonplace quantities. Since

this is the essence of Bayesian updating, Bayesian decision-making is an effective

framework for understanding these effects. First, Griffiths and Tenenbaum (2006)

showed that subjects have reasonable prior probabilities on common everyday

quantities. When asked to predict human life spans or the box-office take of movies,

subjects showed that they had priors similar to the actual distributions of these

values, as would be useful for Bayesian inference. Later, they showed that subjects

also appear to be Bayesian in their ability to use information to make predictions

about the future. In one experiment, subjects were asked for their "gut feeling"

about how often subway trains run when given varying numbers of observations for

when the next train would come. One day they showed up at the station and it was

103 seconds to the next train, another day 61 seconds, etc. In this and similar

experiments, subjects were able to combine multiple sources of information,

combine information with their priors and properly treat independent observations

to make predictions, which is the essence of Bayesian inference (Griffiths &

Tenenbaum, 2011).

Human performance similar to that of an ideal Bayesian subject has also been found in nonverbal, intuitive, domains. For example, Trommershäuser, Maloney, & Sandy (2003) tested subjects in a task that required them to quickly (within 750ms) touch an area on a touch screen. The touch screen displayed a set of penalty and reward areas where if the subject touched them, they would gain or lose the number of points indicated (and at the end of the session points paid out as 25 cents per 1000 points, to make this experiment more easily comparable to existing work on the choosing of lotteries). Because the subjects had to make this decision quickly, their pointing had a certain amount of noise. When the reward and penalty areas overlapped, choosing where to aim one's point became a nontrivial task: aim too close to the center of the reward area and one risks accidentally hitting the penalty area but aim too close to the "safe" edge of the reward area and one risks missing the reward area altogether. Subject performance was well predicted by a Bayesian optimal planner that chose a point of aiming based on the reward/loss function presented on the screen as well as the subject's pointing variability.

In an experiment that can be thought of as a human analogue of Experiment 1 in this thesis, subjects were asked to observe draws from a Bernoulli distribution (essentially, coin flips) whose variable (the bias of the coin) might change without warning (C. Gallistel, Krishan, Liu, Miller, & Latham, 2013). They were asked to estimate the value of the Bernoulli distribution. Their estimates were better modeled by a Bayesian optimal learner than by adjust-as-you-go heuristic models. That was because subjects made decisive and substantial changes in their estimates rather than gradually adjusting to changes.

Of course, just as I pointed out with reference to ecologically-valid heuristics, performing *at the same level* as a given strategy is not sufficient evidence for the claim that subjects are *actually* using that strategy. So, in order to more rigorously test whether subject behavior is consistent with Bayesian models, Maloney and Mamassian (2009) re-analyzed data from the afore-mentioned pointing experiment to examine subjects' ability to transfer information from one task to another. If subjects are performing a Bayesian calculation when performing non-verbal pointing tasks, they are combining a prior function, a likelihood function and a gain function to calculate an approximately optimal response. If instead they are using some fast and frugal heuristic, they will have shortcut one or more of these steps in favor of a simpler solution and therefore will not have learned the prior, likelihood and gain functions separately. They should therefore not be expected to be able to transfer knowledge of these functions to speed up their acquisition of a task that involves only some of them.

The authors noted that in the training phase, subjects engaged in exactly the same task as during the experimental phase except that the reward function was different–though the penalty areas were displayed, they were not yet active so had no effect on the gain function and therefore no effect on subject behavior (subjects learned to point at the center of the reward areas). After several hundred trials, subjects learned to respond within the short time window they were allowed and their motor variability decreased and stabilized. When they moved to the experimental task there was no such learning curve: there was no trend in their aiming points when sessions began nor through the rest of the session. This means

that when confronted with a task with already known prior and likelihood

functions, subjects did not have to re-learn or even readjust their behavior when a

new reward function was introduced. It is difficult to see how subjects would be

able to plug in a new reward function if their behavior was being directed by simple

strategies, be they adaptive heuristics or maladaptive biases.

We have now seen the usefulness of Bayesian models in categorization,

probability change detection and movement planning. Note that these are three

areas that run the gamut from "reason" to "instinct." Categorization has filled

volumes by great minds from Aristotle to Frege and beyond. Detecting probability

changes requires subjects to go from an experience of a stochastic variable over a

number of trials to a verbalized response. Pointing at a high-value area on screen

within 750ms of seeing the display seems to me about as "instinctual" as one can get

since there isn't enough time to deliberate and even the mode of response is

nonverbal. Although choosing a point to aim at in this case is equivalent to choosing

from many possible lotteries (J. Trommershäuser, Maloney, & Landy, 2008), in

which subjects are notoriously prone to error, subjects seem to have no trouble

aiming at approximately optimal points. In all these situations and many more,

subjects perform roughly on par with relevant Bayesian ideal observers. They also

transfer information between tasks and combine information from priors and

independent observations. This means that subjects are essentially doing some kind

of Bayesian analysis in these diverse cases.

The project at hand

Gigerenzer's caution against overfitting and fragile strategies is a valid point but his presentation of algorithms as being either complex and fragile or simple and robust, is overly simplistic. More conceptually sophisticated models can actually be computationally simpler, especially when applied to large and complicated problems. This might be the case, for example, if one is choosing between a set of simple rules (of as-of-yet unknown size) with unknown interactions between them and a sophisticated system that is easily scalable to a large class of problems and may (by nesting) even help you discover the structure of the problem at hand (e.g. Kemp & Tenenbaum, 2008).

The purpose of this thesis is to bolster the case for Bayesian models of decision by demonstrating the usefulness of an Information Processing framework in the animal decision and timing domains. I will establish four basic results that I hope will demonstrate the usefulness of this approach even for decisions under uncertainty by mice:

1. Mice detect changes in probability.

2. The number of trials before this detection is predicted by the information-theoretic Kullback-Leibler divergence.

3. A given subject's timing precision is not constant and can be increased by making the temporal discrimination task more difficult.

4. Subjects distinguish between variability that originates from them and variability that is from the world.

These results serve to demonstrate that mice have the capacity for a necessary component of Bayesian inference (detecting probabilities), are driven by information-theoretic quantities, and can learn and respond to exogenous (as opposed to endogenous) variability. The thinking behind this is that if mice can be shown to approximate Bayesian reasoning then surely it is time for a new, more conceptually sophisticated (which may indeed be computationally simpler) understanding of human decision.

The debate over our notion of rationality has deep intellectual and far-reaching practical implications, as outlined in the suddenly popular books on this topic. Some argue that our uniquely human, careful and deliberate reasoning is actually worse in many situations than our fast and automatic instincts (Gigerenzer, 2007; Gigerenzer & Selten, 2002; Gladwell, 2007, etc.). *Thinking Fast and Slow* (Kahneman, 2011) offers a different basic interpretation of the data: that because our instincts are mere approximations of optimal solutions, we must be even more careful and deliberate in our thinking to reach correct conclusions.

I believe that the work applying Bayesian models to categorization is very encouraging for the prospects of understanding human deliberate decisions. It may, therefore, help to explain deviations from optimality rather than simply cataloguing mistakes. Now, I present what I take to be strong evidence that even decisions under uncertainty by mice are similarly sophisticated and best understood in the same framework: they are sensitive to abstract, information-theoretic quantities and adjust to changes in these quantities abruptly; as if by calculating their behavioral shift and then executing it, rather than gradually tweaking some internal parameter.

I take this to be a step toward understanding rationality, what cognitive capacities we share with animals and therefore by contrast what it means to be human.

General Method

This thesis involves a number of variations on the same basic task, all of which used mice as subjects. The variations are described in the experiments in subsequent chapters but each experiment manipulated some parameter in our interval-timing task, which we refer to as the Switch task. The most important response-variable for all of the experiments was the time at which subjects switch from poking in one hopper to another (the switch latency). Each experiment used this single measure to investigate subjects' models of their environment. I now proceed to describe the general experimental protocol and apparatus. The ways in which each experiment departed from the basic switch task are described in the Methods section of those experiments.

Subjects were first trained to interact with their environment (matching and autoshaping procedures, below) and then tested in one of the variations of the switch task. All tasks were active for two periods per day, one around "dawn" and the other near "dusk" (as determined by the onset and offset of the house lights in their boxes).

Apparatus

The subjects' cages contained three feeding hoppers on the same wall. The two on the ends were designated the "short-side" and "long-side" hoppers and the "control hopper" was in the middle (See Figure *2*). Subjects ran in the 24-hours-per-

day, 7-days-per-week automated testing system and were analyzed by the Time Stamp

analysis system both described in Gallistel *et al.* (2010) and Gallistel *et al.* (2013). Using

these systems, subjects had 24-hour access to their test chamber, which was a Classic

Modular Test Chamber for Mouse from Med Associates (dim: 15.9 cm x 14.0 cm x 12.7

cm) fit with three Pellet Receptacles for use with Pellet Detection, also from Med

Associates. These receptacles were our feeding hoppers, which recorded pokes by the

subject and were the place where pellets were delivered. In the experimental procedures

subjects subsisted entirely from earned pellets, which were 20 mg Purina grain-based

from WF Fisher and Son, Inc. The experiments used different species of mice because

they were piggybacked on other experiments, which were genetic screeners. The species

are indicated in the Methods section of each Experiment. Because of the high level of

performance on our tasks, we have no reason to believe that genetic manipulations

impaired performance.

As mentioned above, subjects lived in the testing environment 24 hours per day,

subsisting entirely off food earned in the experiment, and their actions were cataloged

and time-stamped as they occurred. Analyses were performed in quasi-real time (twice

per day) to keep track of their daily progress in learning the task and adapting to new

conditions.

Training: Matching and Autoshaping

Subjects were first trained on a matching task, in which the feeders

connected to the two feeding hoppers (the long and the short hoppers) were armed

at random intervals so that poking in the hoppers would release a pellet. Next, the

subjects underwent an autoshaping procedure, in which they initiated trials using the control hopper, after which one of two feeding hoppers lit up and a reinforcer was delivered after a set delay. The delay for the left (short) side and right (long) sides were typically easily distinguishable, such as 3s vs. 9s in Experiment 1. This prepared subjects for the primary task: the interval-timing task we call the switch procedure.

Primary task: Switch

In the Switch task, the control hopper was illuminated when the subject could start a trial, which occurred after an exponentially-distributed delay following the end of the previous trial (during the active phases). A poke to the illuminated control hopper extinguished the light there and began a trial. Once a trial began, an unsignaled draw from a Bernoulli distribution determined whether the trial was a "short trial" or a "long trial." The uncertainty about whether the trial would be short or long was an exogenous stochastic variable, which we manipulated in several of the experiments.
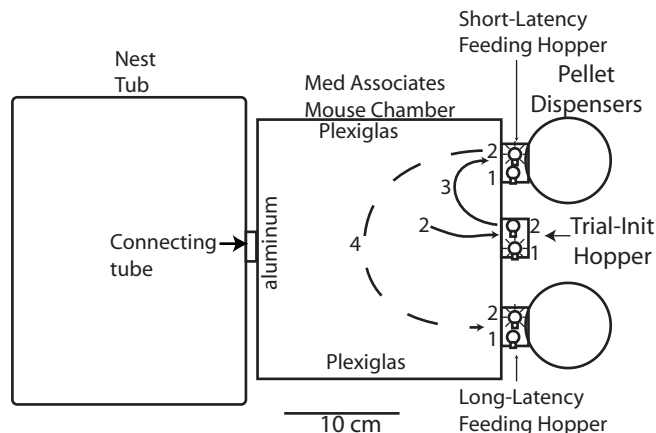
Figure 2. *The experimental environment. In the Switch task, a trial proceeded as follows: 1) Light in the Trial-Initiation Hopper signaled that the mouse could initiate a trial. 2) Mouse approached and poked into the Trial-Initiation Hopper, extinguishing the light there and turning on the lights in the two feeding hoppers (trial onset). 3) Mouse went to the short-latency hopper and poked into it. 4) If, after the short feed latency (e.g. 3 s) had elapsed since trial onset and poking in the short-latency hopper did not deliver a pellet, mouse switched to the long-latency hopper, where it got a pellet there in response to the first poke there after the long feed latency (e.g. 9s had elapsed). Lights in both feeding hoppers extinguished either at pellet delivery or when an erroneously timed poke occurred. Erroneously timed pokes are referred to as "time-outs" because they occurred when the subject poked at the long-side hopper during a short trial after the short-side hopper was armed or when the subject continued to poke at the short-side hoper on a long trial after the long-side hopper was armed. Short trials lasted about the short feed duration and long trials about the long feed duration, whether reinforced or not: If the mouse was poking in the short hopper at the end of a short trial, it got a pellet and the trial ended. If it was poking in the long hopper, it did not get a pellet and the trial ended at the short feed latency. Similarly, long trials ended at approximately the long feed latency: If the mouse was poking in the long hopper, it got a pellet; if in the short hopper, it did not. A switch latency is the latency of the last poke in the short hopper before the mouse switches to the long hopper. Only the switch latencies from long trials were analyzed.*

On a short trial, whether or not a subject was given a pellet was determined by which hopper he poked in first after the short feed latency had elapsed: if this first poke was on the short-side hopper, the subject was reinforced and the trial ended; otherwise the trial simply ended without reinforcement. On a long trial, reward was determined by the first poke after the long feed latency had elapsed: if this first poke was on the long-side hopper, the subject was reinforced; otherwise

the trial ended (Figure *3*). Subjects had no way of knowing which trial type they were in until the trial had ended, but they quickly learned a strategy that allowed them to get most of the possible reinforcers: poke at the short-side hopper until the short feed latency had elapsed and then switch to poking at the long-side hopper (Figure *2*). If the subject had not been reinforced on the short side after the short feed latency has elapsed, the subject knew it was in a long trial so had only to start poking on the long side before the long feed latency has elapsed. The uncertainty about how much time had elapsed since the start of the trial–upon which the decision to switch from the short to the long hopper depends–was the endogenous stochastic variable and the main variable of interest in most of the experiments below.
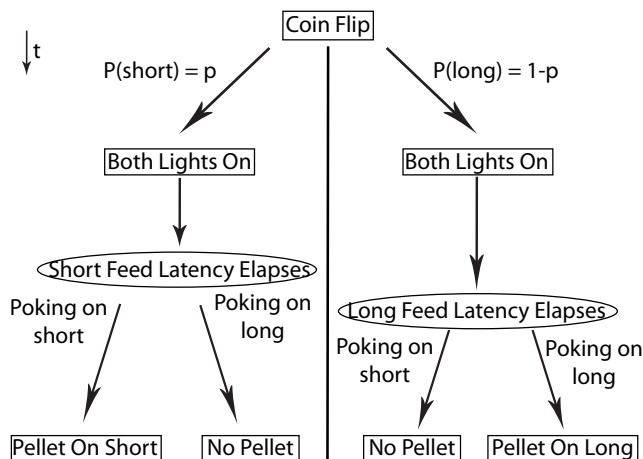

Figure 3. *The basic switch procedure laid out visually. In each trial, a coin of weight p was flipped to determine if it was a long or a short trial. Both hopper lights came on until the first poke after the feed latency elapsed. On a short trial, if the first poke after the short feed latency elapsed was to the short-side hopper, the subject got a pellet. Otherwise, the trial ended without reinforcement. On a long trial, if the first poke after the long feed latency elapsed was to the long-side hopper, the subject got a pellet. Otherwise, the trial ended without reinforcement. Both lights were extinguished at the end of a trial.*

This strategy may be thought of as the subject starting each trial under the assumption that the current trial is a short one. As time passed without being reinforced on the short side, the subject had increasing evidence that the short feed latency had elapsed and therefore that the trial was a long one. The subject could then switch when sufficiently confident that the trial was a long one. This strategy depended crucially on the criterion one sets on one's confidence that the short interval has already elapsed that motivates a switch to the long hopper. Since mice are not perfect timers, whatever decision criterion the subjects used to decide when to switch, their behavior produced a distribution of switch times. So, while switching at any time between the short and long feed latencies would ensure a reinforcer on every trial, aiming to switch at a time very close to short feed latency was a risky proposition: it would be easy to overestimate the time elapsed and switch too early, thereby missing the reinforcer were it a short trial. Similarly aiming too close to the long feed latency would risk switching too late and missing the reinforcer, were it a long trial. Thus, we can think of the subject as engaged in signal detection: attempting to tell whether its own sense that the short feed latency had elapsed was due to the noise inherent in its timing mechanism or if it was veridical. To frame the problem in Bayesian terms, the optimal target switch time was the time at which the posterior probability that the current trial was a long one exceeded the posterior probability that the current trial was a short one.

Switch behavior

Balci (2007) found that the primary component of these switch distributions is normal and centered near the optimal decision criterion given that subject's

timing accuracy (coefficient of variation). Thus, if we interpret the distribution of switch times as the subject's noisy attempts to switch at a specific time, the time they have selected is approximately optimal, as given by the description above. The timing of their switches was centered around t such that t is the intersection of the two weighted normal distributions:

$$p * \mathcal{N}(t, S * CV) = (1 - p) * \mathcal{N}(t, L * CV)$$

where $\mathcal{N}$ is the Gaussian distribution function, p is the probability of a short trial, CV is the subject's coefficient of variation in timing tasks (and thus $S * CV$ is the variance in their timing distribution for the short feed latency), S is the short feed latency and L is the long feed latency (See Figure 4 for an example and Experiment 1 for further discussion).

Note that the optimal target switch time, t, depended on all four of these quantities:

1) If the base rate of short trials, p, were high, it would have been more important that the criterion not cause too many premature switches and thereby miss a large portion of the short trials, which meant a later t (and conversely for low p).

2) As CV increased the variance of the long-trial distribution grew faster than the variance of the short-trial distribution (since $S < L$). The relative increase in the width of the long-trial distribution meant that their intersection moved left, meaning that as CV increased t decreased.

3) The short and long feed latencies had an obvious and direct impact on t: the larger the value of S, the larger t must have been to not miss large

portions of short trials and the smaller the value of L, the lower t must have been to not miss large portions of long trials.
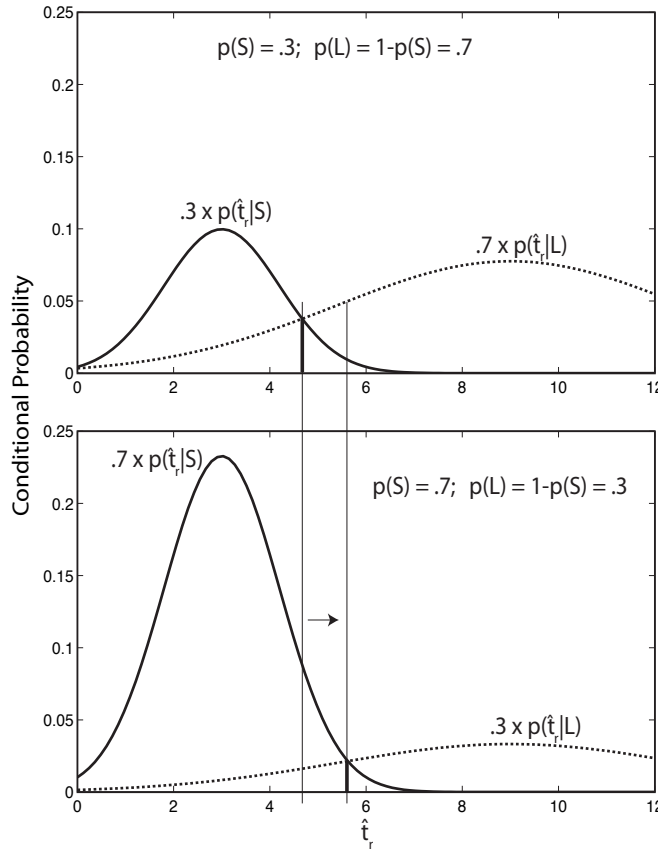

Figure 4. *The conditional probability distributions of the subjective times of reinforcement, $\hat{t}_r$, on short and long trials, $p(\hat{t}_r|S)$ and $p(\hat{t}_r|L)$, scaled by the prior probabilities, $p(S)$ and $p(L) = 1 - p(S)$. The optimal subjective time at which to switch was when the posterior odds were equal (heavy short vertical lines at the points where the distributions intersect). Increasing the prior odds of a short trial (lower panel) shifted this optimum to the right (toward later times). In accord with the well-established scalar variability in subjective elapsed times, the standard deviations of the distributions were proportional to their means.*

Models of the data

As mentioned above, most of the experiments in this dissertation focused on analyses of a single quantity, called the switch latency. This is the amount of time from the start of a long trial until the subject leaves the short hopper for the long hopper. These switch latencies can be seen as samples from a distribution that, as

we have just laid out, depends on many abstract quantities such as the base rate of short trials and the CV of the subject's timing behavior. For the most part, the distribution of switch latencies turned out to be approximately Gaussian but with a small additional component.

This offers a natural interpretation of the observed behavior: the mean of the normal distribution corresponds to the mouse's target switch time (the time at which the mouse, on any given trial, reaches its decision criterion to switch sides) and the standard deviation is a measure of the noise that comes from the mouse's systems of measuring time and enacting their decision to switch once the decision is made.

The addition to the Gaussian distribution was a typically small portion of switches that occurred even before the short latency had expired (see Figure 5 for an example). For the most part, this dissertation focuses on the analysis and interpretation of the Gaussian component of the data. Thankfully, these impulsive switches were typically few in number and easily separated from the Gaussian component. In some experiments, the impulsive switches have been best modeled by an exponential distribution while in others they have been best fit by a Weibull distribution. In either case, none of the results discussed below have depended on such choices of modeling. While these impulsive switches may be of interest to those studying cognitive control, they were only analyzed here so they could be set aside, so as to clarify the analysis of the Gaussian component.
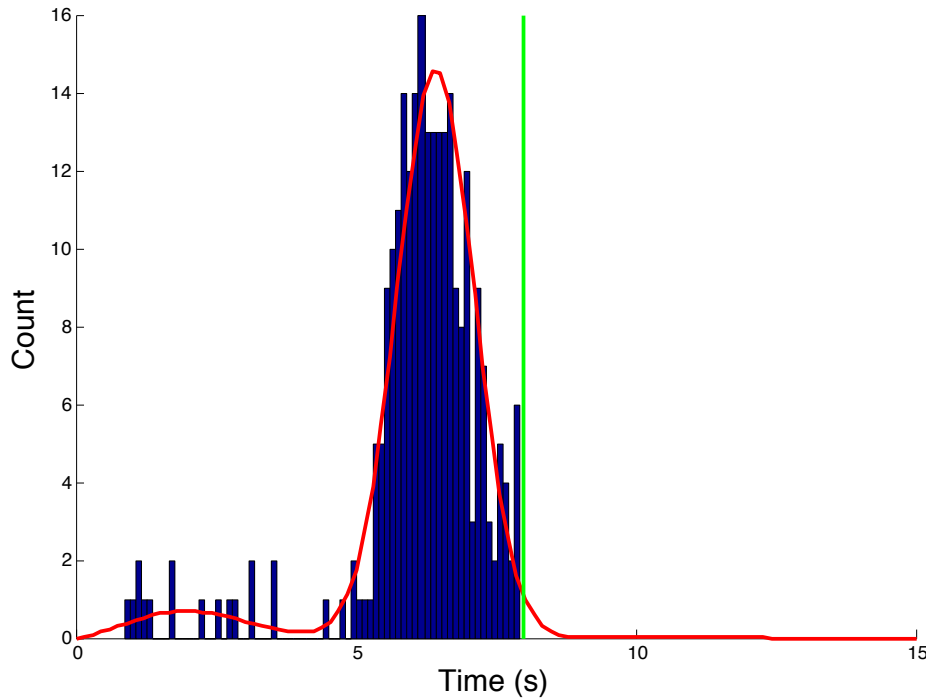
Figure 5. Sample distribution of switch latencies from a single subject in a single session. The data were censored at the long latency (green vertical line) and we see a small portion of switch latencies lower than the short latency, which we call "impulsive" switches. These are discussed further in Experiment 3 but for the most part they were few in number and easily separable from the Gaussian component of the data so were not a large concern for the experiments presented here. Still, they may be of interest to those studying impulsivity and control and warrant further study on their own.

Preview of experimental results

The experiments below expand on Balci (2007) by manipulating these quantities and examining their effect on the distribution of switch times subjects produced. Experiment 1, which was my published masters thesis (Kheifets & Gallistel, 2012b), shows that subjects shifted their behavior in response to a new p soon after it was put in effect and that this behavioral shift was abrupt, rather than gradual. This result greatly constrains the possible models of timing behavior by ruling out optimization techniques that proceed by gradual, trial-and-error adjustment. Experiment 2 examined smaller changes in p and shows that the

number of trials required for subjects to make such a behavioral shift is best understood in terms of the Kullback-Leibler divergence between the p before and the p after. This is a quantity central to information theory but with some counterintuitive properties such as asymmetry, which we see borne out in our data. Experiment 3 manipulated the long feed latency to show that subjects were capable of more precise timing when the task is made more difficult. Finally, Experiment 4 made the short and long feed latencies variable rather than constant. This raises deep theoretical issues about the nature of temporal perception. Most importantly, it shows that subjects treat variability in their environment differently from variability in their temporal perception system: they differentiate between endogenous and exogenous sources of variability.

Experiment 1: Abruptness

Introduction

This experiment is designed to answer one simple but important question: when the probability of an event changes in a mouse's environment, is the resulting change in behavior abrupt or is it gradual? As I will argue below, subjects' performance on the switch task speaks volumes about their ability to represent and manipulate abstract information, rather than use simple heuristics, to calculate approximately optimal responses. I will spend a great deal of time setting up the mathematical background of this experiment and discussing the implications of the present data because subsequent experiments will build on this one. This experiment and the interpretation of its results lay the foundation for the rest of the dissertation.

Adapting to changing probabilities in one's environment is a ubiquitous problem for all animals. This is an especially difficult task in comparison to, say, adapting to a change in temperature because probabilities are not directly observable. It is only through repeated observation of similar events (several flips of the same coin) that we can estimate hidden variables such as the probability of a heads. Estimating probability, however, is just the beginning of the problem facing animals both in the wild and in our task.

Factors influencing the probability of an event are constantly changing, so that one must also update one's probability estimate. So, whenever an animal

observes a string of surprising events it must infer whether its model of the world is outdated or if this is just random stochastic variation. Doubtless, animals successfully navigate the world with its many, changing, hidden variables. I argue that their performance in detecting and responding to changes in the values of a hidden variable cannot be explained by reinforcement models which gradually tune behavior to maximize reward. Rather, they seem to detect changes and abruptly change their behavior to align with a new estimate of the hidden variable. This is like a probabilistic analogue of the results in the matching domain from Gallistel, Mark, King, & Latham (2001), in which rats responded to sudden changes in their rate of reward suddenly and about as quickly as an ideal Bayesian detector.

In the switch task, optimal performance requires estimating the probability of each trial type and noticing when that probability changes. In fact, each response from an optimal agent can be seen as an estimate of what that optimal agent takes the probability of a given trial type to be. Without signaling subjects, I changed the probabilities of the trial types and then quantified the response to this change. One can see from a scatter of the raw data (Figure 6 below), that subjects responded to changes in the probability of trial type shortly after the change occurred. The bulk of the analyses below consist in quantifying what is clear in Figure 6 below–the distribution of switch latencies shifts soon after a change in probability and the shifts are abrupt.

The changes from the old to the new distributions are also abrupt, making them indistinguishable from step changes. This suggests the explicit detection of the change in probability, followed by the computation of a new decision criterion (a

new target switch time), which requires an enduring representation of the subject's temporal uncertainty together with a new estimate of the probability of a short trial. The abruptness of the adjustments does not appear to be consistent with the gradual attainment of a new dynamic equilibrium through "hill-climbing," as in simple reinforcement-learning models. To achieve the observed degree of abruptness, the learning rate parameter must be set very high, but then a reinforcement-learning model would track the stochastic noise in the sequence of short and long trials, which the mice do not do.

The current study deals with the switch timing task which, to review, is influenced by a hidden Bernoulli process: each trial is a "short" or "long" trial as determined by an unsignaled flip of a coin of weight p. On short trials, subjects are rewarded for not acting too early and on long trials, they are rewarded for not acting too late. This means that the weight of the coin has an impact on the ideal temporal decision criterion for an ideal subject: the higher the probability of a short trial, the later an ideal subject with limited timing precision should set its temporal decision criterion. If short trials are more likely, it is important to minimize the risk of acting too early even at the expense of increasing the probability of acting too late, since acting too late is not penalized on short trials. This calculation can be understood in the framework of signal detection theory, which is a branch of information processing.

Two separate levels of signal detection

1. One type of signal detection occurs at the trial-level. On a straightforward view of how a rational agent would perform this task, each trial would begin with

the agent taking the probability of a short trial to be the recently observed proportion of short trials. As time elapsed and the agent waited at the short-side hopper, the agent's confidence would climb that three seconds had elapsed since the beginning of the trial. Since three seconds elapsing without being fed and the current trial being a short one are mutually exclusive, each moment would be increasingly strong evidence that the current trial was not a short one. A rule about when to switch sides easily follows (which the reader will recognize from the General Method discussion): switch sides iff.

$$p * \mathcal{N}(t, S, S * CV) = (1 - p) * \mathcal{N}(t, L, L * CV)$$

where $\mathcal{N}$ is the Gaussian distribution function, p is the probability of a short trial, CV is the subject's coefficient of variation in timing tasks (and thus $S * CV$ is the variance in their timing distribution for the short feed latency), S is the short feed latency and L is the long feed latency. This rule essentially says "when the base rate of a long trial times the evidence that this is a long trial exceeds the same quantity for short trials, switch sides." This is a normative strategy for this task: the switches will be close to the moment at which it becomes more likely that the current trial is a long one, thereby maximizing the number of correct trials.

Again, if the subject knew the exact trial time with certainty this would be trivial: the Gaussian distributions above would instead be Dirac delta functions, which evaluated to 1 at L and S respectively. The subject would then switch at exactly L seconds and never risk missing any trials. Because subjects are not perfect timers, this is instead a signal detection task: at each moment the subject's representation of how much time has past is in part due to the signal (the

representation of time elapsed does go up with actual time) and in some part due to noise. A rational agent must therefore attempt to tell whether its own sense that the short latency has elapsed is due to the noise inherent in its timing mechanism or if it is because it has actually elapsed.

2. A second signal detection problem occurs on the inter-trial level: is the recent sequence of trials drawn from the same distribution that has been generating the trial sequences thus far or has the distribution changed? By the end of each trial, the subject knows whether that trial was a long or a short one: either by being fed or by timing out. That still leaves open the question of what Bernoulli distribution that trial was drawn from. If the subject has seen a lot of long trials and then sees a streak of short trials, this could either be an unlikely set of draws from the same distribution or draws from a new distribution.

How convincing must a streak be in order to be understood as a change in distribution[1]? If an agent wrongly labels streaks as genuine changes, their switch behavior is likely to be erratic; tracking the noise in their stimulus. If an agent wrongly dismisses genuine changes as streaks, they will be slow to adjust to changes and therefore will miss trials due to operating with bad estimates of the base rate of short trials. In order to constrain the possible ways in which subjects might be performing this signal detection, I examine a) how rapidly mice detect a substantial change in the base rate and b) how abruptly they adjust to this change.

I found that the distribution of switch latencies (the number of seconds from the beginning of a trial to the time when they switch from poking on the short side

---

[1] We assume subjects place equal utility on false positive and false negative.

to the long side) depended strongly on the probability of a long versus a short trial, as found by Balci *et al.* (2009), and this dependence was as the signal detection framework would suggest: An increase in the probability of a short trial translated the distribution of switch latencies in the direction of longer latencies, and vice versa.

Method

Subjects were six female C57/BL6 mice from Harlan Laboratories, aged approximately 8 weeks. They were run in the 24-hours-per-day, 6-days-per-week automated testing system and were analyzed by the Time Stamp analysis system both described in Gallistel *et al.* (2010), trained using the matching and autoshaping procedures described in the General Method section. The active periods for the switch procedures were 9pm-11pm and 4am-8am, in which subjects performed the switch task with short feed latency 3s, long feed latency 9s and the probability of a short trial set to .1, .3, .5, .7 or .9 depending on the session. The abruptness with which subjects shifted their timing behavior in response to a new probability value and the latency of this shift were the measures of primary interest in this experiment.

Results

Depending on p, the optimal switch point for a mouse changed: if long trials were more likely, not switching too late became more important for reward maximization than the danger of switching too early. So, when long trials became more frequent, the optimal switch time occurred earlier, which can be seen in

Figure 6: the scatter plot of the subjects' responses (black circles) responds to the change in the probability of trial type (density of red crosses above vs. below) shortly after the trial at which the change occurs (black dashed line). It is worth noting that the majority of the analyses in this paper will be aimed at making more rigorous what is visible to the naked eye in this plot of the raw data: the behavioral shift in response to the change in the stimulus comes shortly after the change and is abrupt.
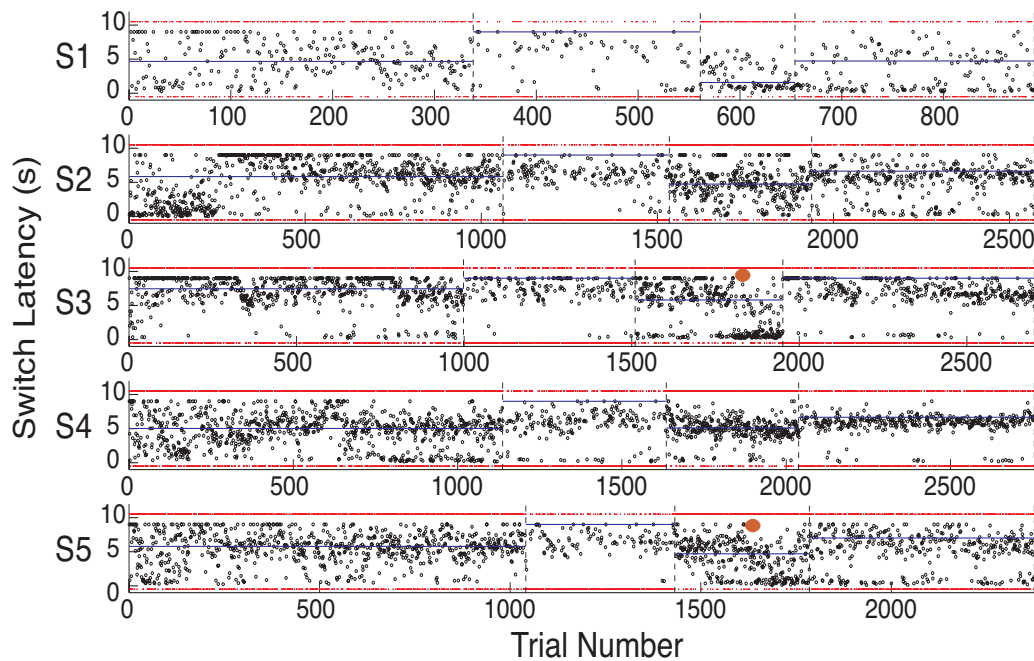

Figure 6. Scatter plot of switch latencies for five subjects. Each black dot represents the latency of a switch on a long trial and the blue lines represent the median of these latencies in a given session. The red crosses that appear in lines above and below the black dots mark long and short trials, respectively. The vertical lines mark session boundaries at which the relative frequencies of the short and long trials changed. Note the change in density of the lines of red crosses that occur at these boundaries. Note further that when the density increases on the top line (and decreases on the bottom), the distribution of black dots shifts downward, away from the denser red crosses. Similarly, when the density of red crosses increases in the bottom line and decreases in the top line, the distribution of black dots shifts upward, again, away from the denser line of red crosses. The seemingly unprovoked changes marked by large orange dots were caused by unilateral feeder malfunctions.
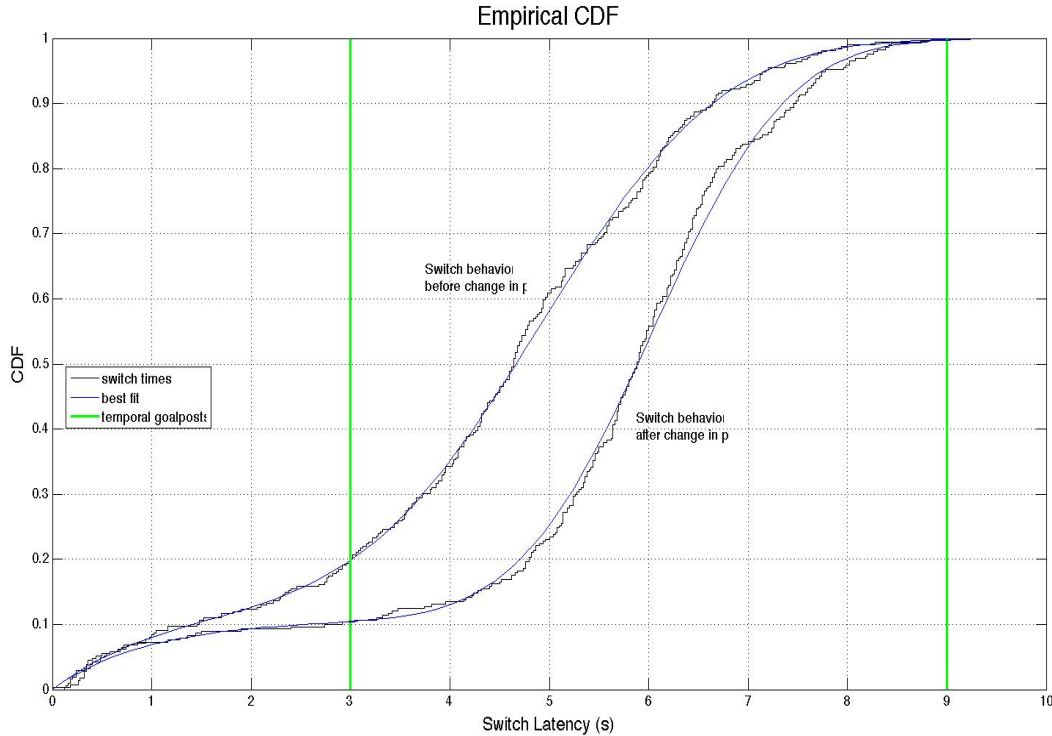
Figure 7. The cumulative distribution of switch times from a single subject under two different probabilities of a long trial (.9 on the left and .1 on the right). All switches between the "temporal goalposts" (3s and 9s) are reinforced, but one sees a strong shift toward shorter wait times when the probability of needing to switch (probability of a long trial) is high.

I found that mice shifted their target switch time in accordance with the change in the optimal switch time (one sees a distribution shift to the left when the probability of a long trial increases, as in Figure 7) and that this change was both abrupt and followed shortly after the change in p. First, I replicated the findings of Balci *et al.* (2009), in that the subjects showed statistically significant differences in the means of the distributions of their switch times when the probability of a short trial was manipulated.

Also in line with the findings of Balci *et al.* (2009), I found that subjects missed very few of their possible reinforcers. Each plot in Figure 8 shows two functions of a subject's decision criterion and their timing precision. First, it shows

the pair of values that best describe the subject's behavior in that session (these are the roughly concentric circles). Second, it shows the proportion of total rewards possible that would be received for each pair of values from a hypothetical subject that went through the same sequence of trials (these are the upside-down u curves). In order to get above 99% of the possible rewards, a subject must have a mean-CV pair that falls below the lowest red curve. Note that as a hypothetical subject's timing becomes more precise (points closer to the x-axis) the subject can freely vary its decision criterion without greatly changing its expected reward. In the degenerate case, when precision is perfect, subjects could put their decision criterion anywhere in the three-to-nine second range and still get all of the rewards. Conversely, the worse a subject's timing precision, the narrower the range of mean values that will produce a given level of proficiency.

Despite the fact that many subjects exhibit temporal precision that would allow fairly large variability while still staying within the range of values that would garner 95% of all possible rewards, subjects tend to have decision criteria close to the mode of the expected value function at their level of precision (the x-value for the top of the closest upside-down u function). This suggests that they are able to perform at near-optimal levels, even when doing so has little impact on their net reward–a calculation that requires sensitivity to one's own uncertainty/variability.

A subject's sensitivity to its own variability has not been well explored in the animal literature but has been demonstrated in humans. In a series of experiments, subjects were asked to quickly hit a goal area on a touch screen while avoiding penalty areas. Subjects successfully chose near-optimal targets, which depended on

the their motor variability: they did not go after small targets surrounded by large penalties when doing so was beyond their pointing precision (Behrens, Woolrich, Walton, & Rushworth, 2007; J. Trommershäuser *et al.*, 2008; J. Trommershäuser *et al.*, 2003).

Finally, there was virtually no overlap between 95% confidence intervals under different short-trial probabilities. This underscores that the behavioral shifts were substantial with respect to our confidence intervals on mouse behavior. Subjects made large shifts in their switch behavior and they did so abruptly.
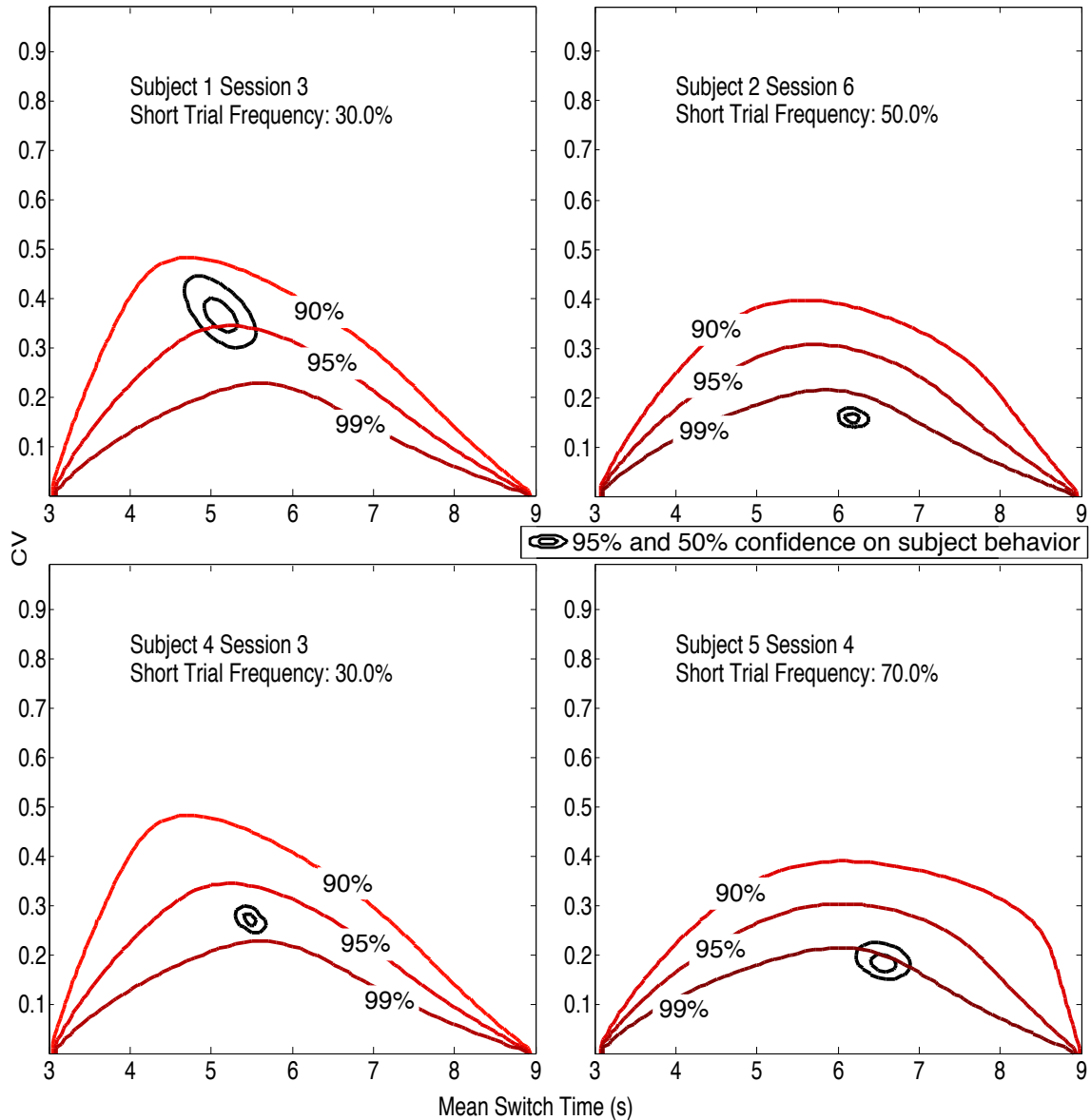
Figure 8. Each subject's behavior (black concentric circles) falls near the expected reward contour (upside-down u-shaped curves) that yields 99% of the total rewards possible in the task. That is to say, the subjects were reinforced 99% of the times that they were given the opportunity to be. Such ceiling-level performance was typical on all the transitions. Assuming that subjects cannot influence their timing variability, the y-position of their behavior is fixed and they merely select the x-component.

Also similar to Balci *et al.* (2009), I typically found between 5-10% impulsive switches in these data (trials in which subjects switched sides even before the short latency had expired). This was a significantly higher rate than could be explained by inaccuracies in a subject's internal timer. I interpret this as meaning, for a small

portion of trials, subjects were simply not doing the task but rather switching sides immediately[2]. This will be discussed in later experiments but in general I do not see this as any more cause for concern than throwing out trials in a psychophysical task when undergraduate subjects are suspected to be not on-task. These impulsive switches are easily separated from the Gaussian portion of the data and analyses focus on this Gaussian component. The purpose of identifying these trials is simply so they do not muddy estimates of the Gaussian parameters.

Returning to the main purpose of this experiment, the question of "how quickly do the mice react to the change in p?" splits up into two more specific questions: 1. how many trials does the mouse need to detect the change in the p value? and 2. Once the change begins, how abrupt or gradual is it? The answer to the first question speaks to how good the mice are at detecting changes in probability; the second speaks to the method by which they change their decision criterion: gradually trial-by-trial or determining a new target value and abruptly switching to it.

In order to answer these questions, I considered two-parameter (corresponding, roughly, to location and abruptness) models of the change. I considered several forms for the model of the transition (Gaussian, Weibull and linear–see Figure 9) but they were all essentially similar: we know the animal's behavior at the beginning and we know it at the end, the question is just how the behavior transitions from one pattern to the other. The different forms of functions

---

[2] Like the trials in which subjects failed to do the timing task by simply switching immediately, there were also a few sessions with slightly more timeout trials than could be accounted for by stochastic variability. These were taken to be trials at which the mouse was "asleep at the wheel" and these were excluded. These were too few in number, however, to materially affect the analyses discussed here.

yielded similar results so I will discuss only the linear class of model for simplicity's

sake.
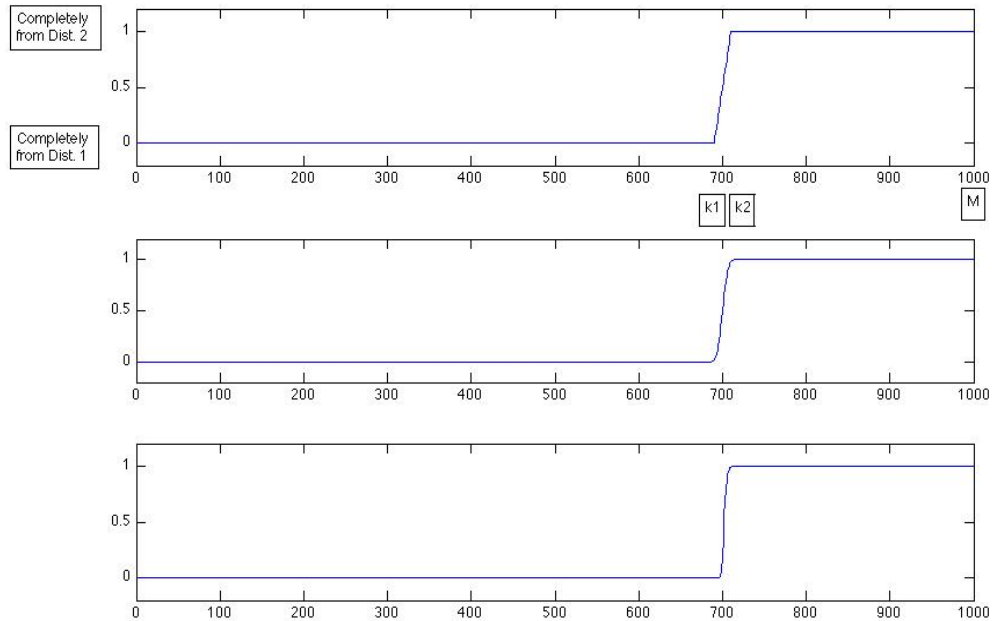
Figure 9. In the range of values plausible for these analyses, the form of the transition function had negligible impact in the range of parameter values that were plausible. The transitions were too arupt to discriminate between the forms tested: linear (top) Gaussian (middle) and Weibul (bottom).

Linear	$$L(x) = \left\lceil \left\lfloor \frac{start - x}{duration}, 0 \right\rfloor, 1 \right\rceil$$
Cumulative Gaussian	$G(x) = \phi(x|start + dur/2, dur * 2)$
Cumulative Weibull	$W(x) = start + W(x|start, dur)$

The linear model states that, given trials 1 to M (in which p-value has

changed somewhere) the animal's behavior is drawn from exponential-Gaussian

distribution1 for trials 1 through k1 and the animal's behavior is drawn from

exponential-Gaussian distribtion2 for trials k2 through M. The behavior between k1

and k2 is drawn from an exponential-Gaussian distribution whose parameters change linearly from those of distribution1 to those of distribution2.

The majority of p-value changes show the same result: the animal's behavior seems completely shifted to distribution2 by the time one has just a few measurements (switch latencies on long trials) from that animal. Looking at a contour plot of the likelihoods for various pairs of location and gradualness, one can see that both values are low, meaning that the change is centered shortly after the change in stimulus and is completed rapidly. Figure 10 shows the marginal distributions of such a plot.

Indeed, across subjects and sessions, there is variation in how wide the confidence intervals are (this is often tighter when we have a high percentage of long trials because we can only take measurements at long trials, which means more data with which to constrain the model in the low range of values) but typically the peak of the marginal likelihood functions are close to zero (in the case of the start of the transition, it is often too low to be taken at face-value for a priori reasons). This means that the most likely model for the data observed is one in which subjects begin the change in their behavior shortly after the actual change and complete the change in their behavior in a small number of trials.
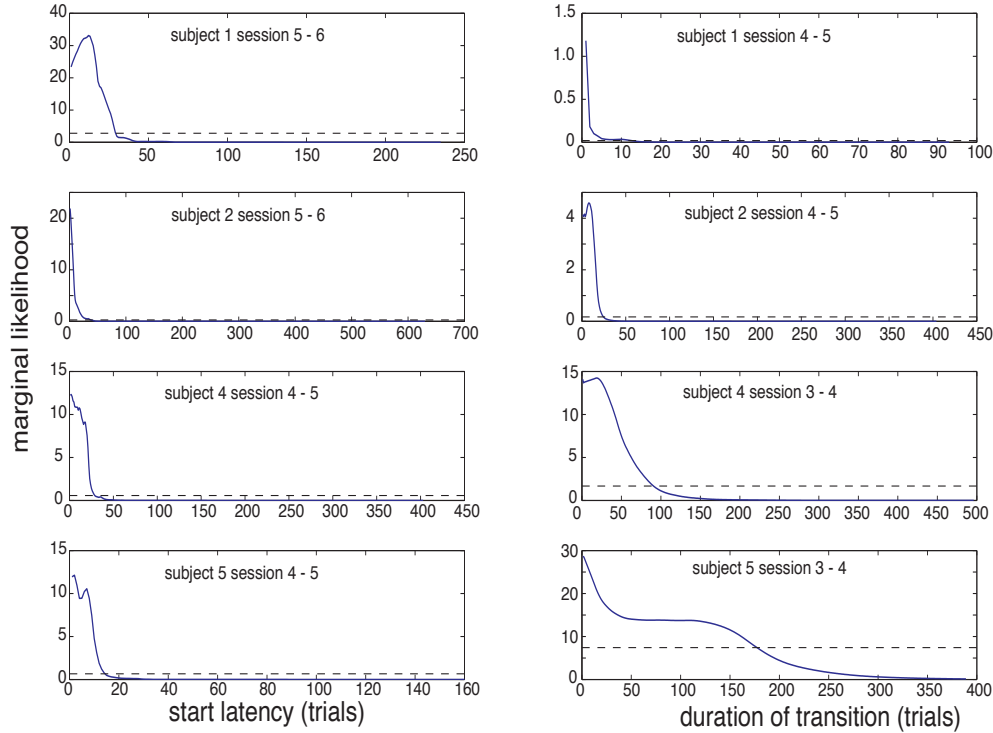
Figure 10. The marginal distributions on the location likelihood (number of trials after actual change in stimulus probability at which linear transition begins) and duration of the behavioral shift (span of trials from start to end of linear transition function) for a sample of eight transitions.

Looking at the marginal likelihood of various transition durations, one sees that the log-likelihood at zero is higher than typical likelihood for other transition durations. This suggests that a simpler, one-parameter, model actually does a better job of explaining this data: the transition is a single step and the only free parameter is where that step occurs. We test this claim below and the Bayes factors of this test can be seen in Figure 14.

All these analyses were also conducted using a 10-parameter model: one in which the four parameters for the expgauss distribution before the transition, the four for distribution after the transition and the location and abruptness of the transition were all allowed to vary. I used Metropolis-Hastings MCMC to sample

from the likelihood function of this 10-parameter model and found that the maximally likely values for the start and duration of the transition were the same in the 10-parameter model as in the 2-parameter model – integrating out the other 8 parameters which represented the distributions before and after the change, did not change the overall shape of the marginal likelihood functions for the two parameters of interest[3]. The likelihood functions derived from the 2-parameter model were essentially the same as those obtained by applying the 10-parameter model and averaging out the other 8 parameters.

Importantly, my basic conclusions do not rest on particular method of examining the likelihood function. Whether one looks at the maxima of the marginal distributions, their expectations, etc. one reaches the conclusions that behavioral shifts occur soon after the change in probability and are completed within a small number of trials. This is illustrated in Figure 11, where one can see that over half the behavioral shifts have a duration of approximately 20 trials or less and start approximately 10 trials after the probability change.

---

[3] I also confirmed that the maximally likely exponential-Gaussian distributions obtained by fitting the 10-parameter to the switch behavior before and after the change were the same as the exponential-Gaussian distributions we obtained by fitting each data set individually.
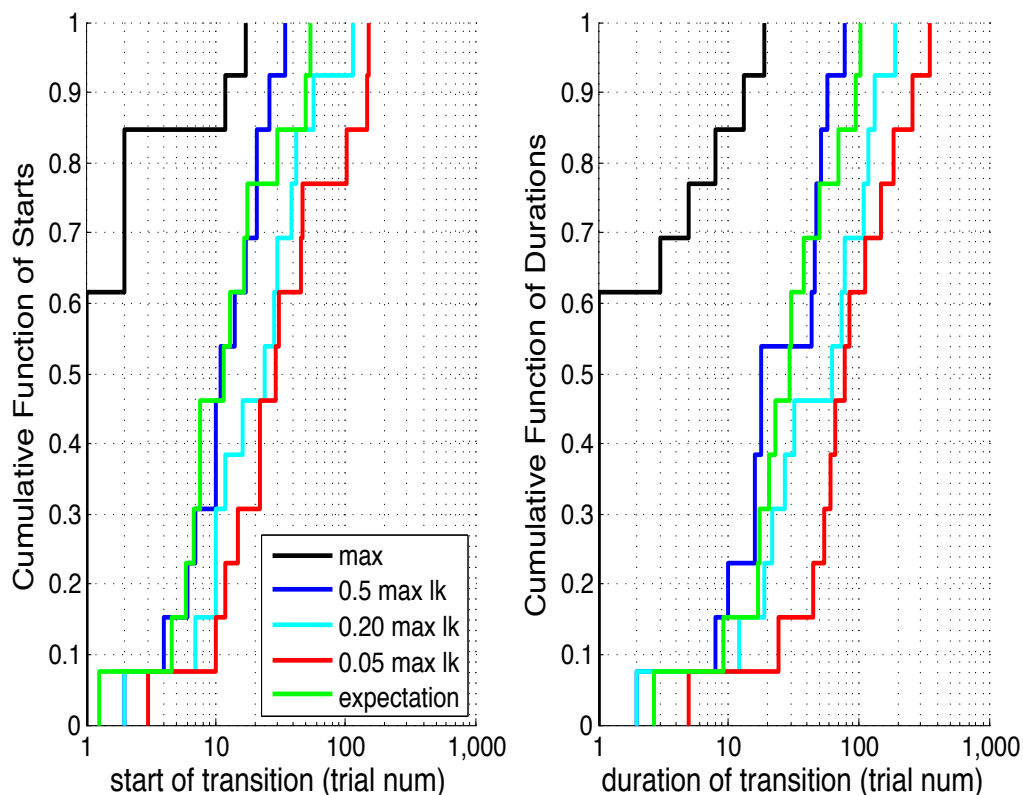
Figure 11. Cumulative distribution of locations (number of trials after actual change in probability of a long trial) and durations of the behavioral shifts. The CDFs are shown using various criteria for the cut-off point: Maximally likely start/duration, the lowest start/duration with at least ½ the maximum likelihoods, 1/5 maximum, 1/20 maximum and the expectation of the likelihood distribution. The ½ max and expectation criteria seem to best summarize the data for our purposes, but none of our analyses hinge on chosing one criterion over the others.

Discussion

One can imagine proficient performance on the switch task from two basic types of machines: machines that use trial-and-error to find viable target switch times and machines that calculate near-optimal switch criteria given the data available to them. These data suggest that subjects engaged in the second type of behavior because of four basic empirical findings about the shifts in their switch latencies:

1. Behavioral shifts occur shortly after the stimulus probability change (as soon as there is objective evidence that there is been a change, see Figure 12).

2. Behavioral shifts are completed in few trials; many shifts are completed so rapidly that the transition can't be distinguished from a step.

3. Behavioral shifts are accurate. They are often closer to the optimal value than could be expected to be found with the limited amount of feedback subjects received from missed pellets.

4. Behavioral shifts are based on detecting changes.

I will now elaborate on each of these findings.

Shifts occur shortly after changes

Speaking to the first point, we have seen that the marginal likelihood distributions are skewed in favor of shifts that begin shortly after the objective changes in the probability of a long trial. This shift latency, however, is difficult to interpret in isolation. Is 10 trials until the behavioral shift a lot or a little? Does that mean that subjects are quick at picking up changes in p or slow? These questions demand that our data be put in context and an Information Processing framework provides us with ideal observers to provide this context. To answer these questions, we can examine, given the particular sequence of long and short trials a subject saw on a given transition, the objective odds that a change has occurred at the trial that subject began its shift in behavior.

Figure 12 plots the objective odds that a probability change had occurred at the trial when subjects began to shift their behavior. This is plotted for various

criteria for determining at which trial the behavioral shift began to emphasize again that the value of this criterion does not greatly impact our interpretation. We see that over half of changes begin by the trial at which the objective odds reach 3:1. This supports the assertion that behavioral shifts begin nearly as soon as there is substantial evidence that a change has occurred–that learning occurs quickly when trials are surprising. This is in line with the Bayesian re-analysis of classical behavioral data by Courville, Daw, & Touretzky (2006) which showed that in a variety of classical tasks such as Pavlovian conditioning, experiencing surprising events (defined in the objective, statistical sense) predicted the acquisition of the change in task parameters.
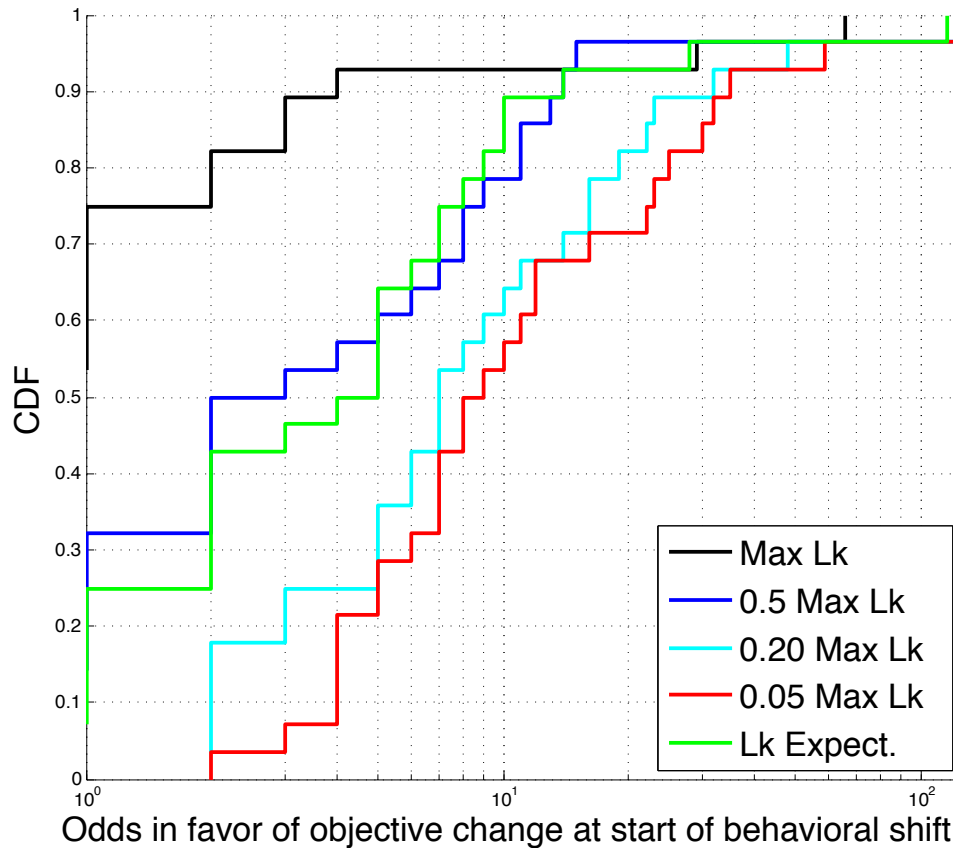
Figure 12. The cumulative distribution of the objective odds that a change had occurred as of the estimated start of a transition. The different colors are for different estimates of the start (maximum likelihood, expectation, and confidence limits of increasing stringency--see legend). Expectation and ½ maximum likelihood seem to best capture the central tendency of the data but again, the basic interpretation does not depend on particular choices of this criterion.

Note that a significant number of shifts begin before there is even strong evidence that p has changed (approximately 30% of shifts begin at or before the odds of a change are less than or equal to 1:1). Keep in mind, however, that the estimates at the low end of the scale are especially volatile because of the confluence of two factors: the model being poorly constrained in the low shift latency values (where it is most likely) and the abruptness with which odds rise after a change.

In some cases, the estimates of the start of the transitional shift are so low (and therefore unconstrained by the data in that range) that the best estimate for

the start of the change are before substantial evidence has accrued that the change had occurred. Because of the rapid rate at which the odds grow, estimating the start of the transition even a small number of trials early can cause the odds at the start to be very low indeed. Likewise, estimating a few trials late can cause the odds to be very high, which is why we see the starts of transitions with sometimes well over 10^4 odds in favor of there having been a change. See Figure 13 for examples.
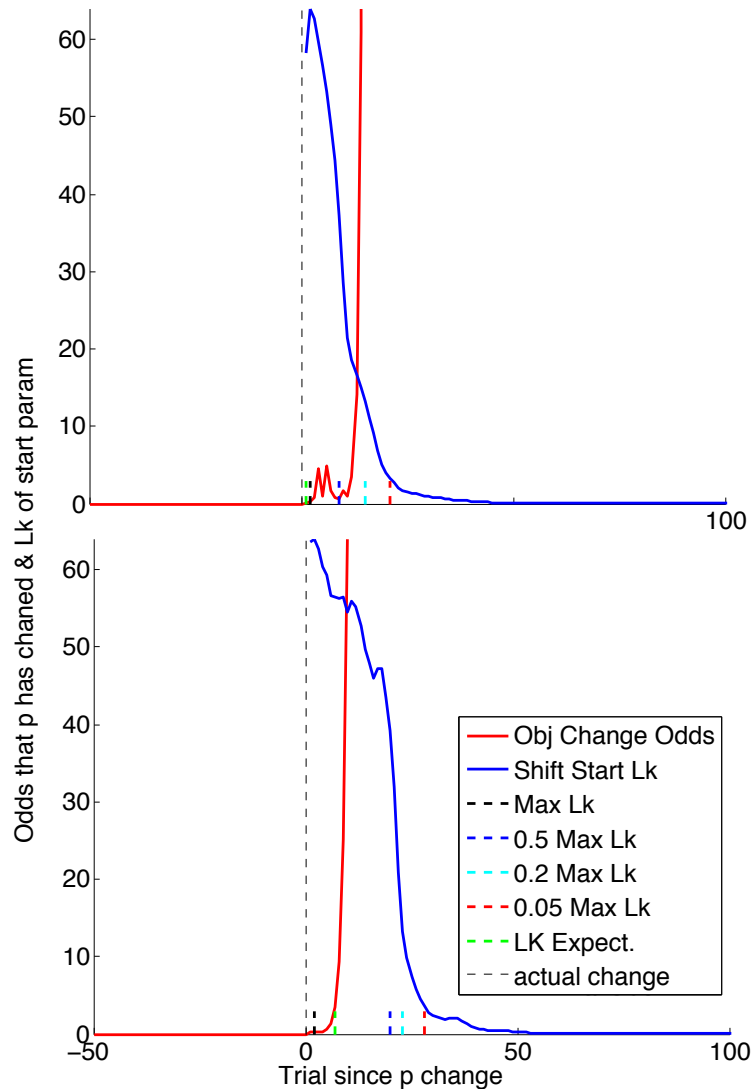
Figure 13. The likelihood function for the location of the start of the behavioral shift superimposed with the objective odds that a change in probability occurred (based on the sequence of trials leading up to the current one). Estimates for the start of the behavioral shift are in dashed lines for various decision criteria: ½ maximum likelihood (blue), ⅕ maximum likelihood (cyan), $1/_{20}$ maximum likelihood (red), and the expectation of the likelihood function (green). Note that the abruptness of the rise in the odds that an objective change has occurred means that we see high variability in the odds when a behavioral shift is estimated to start. This is especially true when the location of the start is in the neighborhood of the sharp increase in odds.

Shifts were abrupt

Once behavioral shifts begin, they are completed quickly. The marginal

likelihood functions for the gradualness parameter favor shifts so abrupt as to

nearly be indistinguishable from a step. If we compare the hypothesis that subjects make step-like shifts to the hypothesis that they made gradual shifts (uniform prior over the range of plausible gradualnesses) we see that the Bayes factor favors a step change 10:1 for approximately 75% of changes.
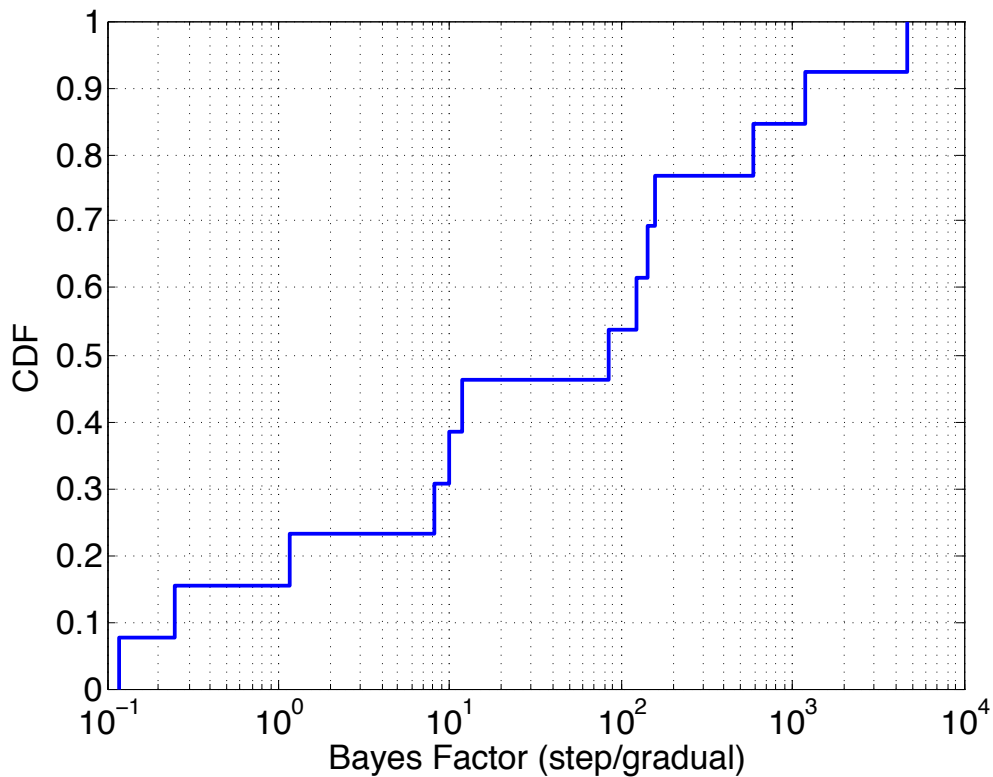
Figure 14. Cumulative distribution of the odds on the null (values larger than 1 favor a step change, rather than a gradual one). The odds favor of a step change for 75% of transitions, with over 50% of transitions favored over 100:1.

Speaking to the larger point of this thesis, these results support a model of subject behavior that explicitly represents the probability of a long vs. a short trial and a subject's own timing precision. We see in Figure 7 that approximately optimal selection of one's decision criterion depends both on the frequency of long vs. short trials and on the subject's timing precision. The fact that subjects' observed temporal decision criterion was close to the optimal for their level of timing precision suggests that their decision criterion is set as a function of these variables

(probability of a short trial and subject timing variability). This means that the variables must be represented by the subject.

Shifts Not Based on Differential Feedback

Fundamentally, reinforcement models base behavior on the subject's level of reinforcement. Such models cannot account for the change in behavior we observe for two reasons. First, for many changes, subjects shifted their behavior with little difference in their level of reinforcement. That is to say, the rate of pellet loss shortly after the change in the probability of a long trial (solid lines in Figure 16), when subjects are adapting to the change in their stimulus probability, is nearly indistinguishable from the rate of pellet loss leading up to that change (dashed lines in Figure 16). Reinforcement learning models therefore have no difference in the level of reinforcement received by the subject to motivate the observed shift in behavior.
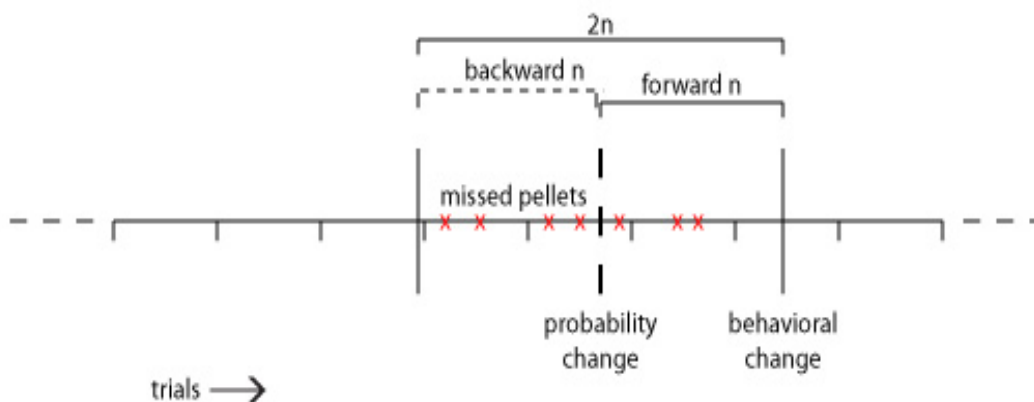
Figure 15. The number of trials to the midpoint of the behavioral shift, n, is determined by the criterion used for the start and duration of the shift as in Figure 11. One can then look at the n trials after the change in probability to find the number of reinforcers missed while the subject adjusts to the new probability of a long trial. We can then compare this number with the number of reinforcers missed in the n trials leading up to the change in probability (when the subject is already adjusted to the previous value) to see if that number is typically different from the number missed while the subject gathers evidence that a change has occurred and shifts its behavior accordingly. These two values are compared in Figure 16 for varying criteria on the start and duration of the behavioral shift.

Keep in mind that their behavior does shift and this is in response to a change in probability: it becomes closer to optimal given their level of timing precision, as can bee seen in Figure 8. The level of reinforcement, however, does not drive the behavioral shift. If we imagine the expected gain function for a given probability value as a broad hill sitting on top of the mean-CV plane, when the probability changes, the hill jumps to a new location. The new hill is also rather broad, so the old position of the subject, near the top of the old hill, is not much lower on the new hill. Still, they shift their behavior so as to stay on top of the new hill. This analysis shows that even when subjects are getting nearly every reward possible, they are able to adjust their behavior to stay comfortably on top of the hill of expected rewards.

Second, as Figure 16 shows, in 30% of cases the subjects were half way through their behavioral change before a single reinforcement was missed. Since these shifts in behavior preceded even a one-pellet difference in the level of reward, the level of reward could not contain any information about the existence of changes that would then drive the shifts in behavior. Adjust-as-you-go heuristic methods of learning cannot explain learning in these circumstances because the adjustments are made based on feedback from the environment. Parameter tuning by trial-and-error cannot account for the present data because subjects had so few trials that would inform their tuning before their behavior changed–keeping in mind that they do not switch at all on short trials and so get no feedback on the appropriateness of their target switch time.

On the basis of these data, a transition is typically half complete after 15 trials (fully complete after 20) (Figure 11). In that number of trials, subjects are likely to have missed only 2-3 pellets (Figure 16). Most impressively, in many cases the transition was half complete before the subject missed a single reinforcer. This is a knock-down argument against learning by differential reinforcement in these cases because no reinforcements have been missed.
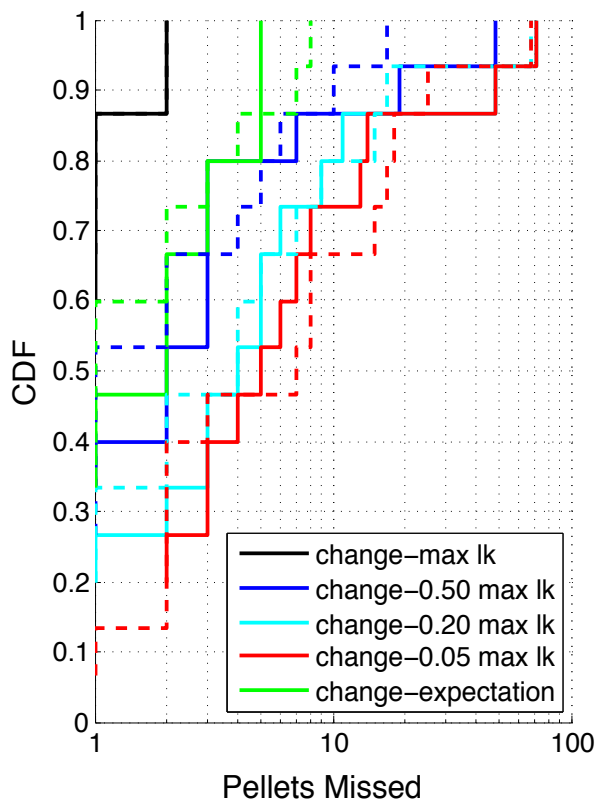
Figure 16. Cumulative distribution of the number of pellets missed after the probability value changed but before the animal was half way through its behavioral shift, using various likelihood criteria for the start and duration of the behavioral shift. Note that typically very few pellets were missed before the subject shifted its behavior: in a sizable number of cases, not a single pellet was missed. Moreover, the number of pellets missed over this period (solid lines) is almost indistinguishable from the number missed during the (same number of) trials before the shift (dashed lines). This underscores how quickly behavioral shifts are made relative to the feedback the subject receives about the consequences of its behavior.

Shifts Are Based on Detecting Changes

If heuristics that rely on the level of reinforcement to drive behavior cannot explain the present data, subjects must be using models that explicitly represent the task parameters instead. Dayan (2012) makes a similar distinction between model-based control as opposed to model-free control. Dayan characterizes this difference as how much of behavior is driven by inference as opposed to experience. In our

case, it is perhaps more useful to characterize the distinction as how much of the world must be represented by an agent.

We have shown that an agent cannot exhibit the behavior we have observed by merely doing calculations on changes in its rate of reward; it must represent some estimate of the probability of a long vs. a short trial. I argue that the quick and decisive behavioral shift we see is the result of a model-based control process that represents not only the probability of a long vs. a short trial but also the changes in this value. This is as opposed to simpler models that adjust to changes in probability quickly by using an estimate of long trial frequency that is only impacted by the last few trials (discussed below). This argument underscores the abstract and complex calculations that underlie subjects' behavior.

We know that subjects must use an estimate of the probability of a long trial and that getting an estimate of probability from experience requires some sort of averaging. If this averaging places significant weight on trials in the distant past, the model will not be able to respond to a change in probability as quickly as subjects did. So, I checked, for various small memory sizes m, whether my results could be simulated by a simple strategy that based switch time off of the frequency of long trials in the last m. These simple strategies are not reinforcement strategies because they are based on an estimate of the probability of a long trial but they do not explicitly represent a change in this probability. Rather, they adjust to a change quickly by not allowing anything more than a few trials back to influence behavior. Thus, the change in the stimulus causes the behavioral shift without the subject's having to represent that a change occurred.

To test whether such a strategy could explain the present data, I sorted all switches by the frequency of long trials preceding them and then compared, via Bayes factor, whether the switches seemed to be drawn from the same distribution or from different distributions. I did this for several sizes of memory (1, 3, 5) with little difference in the results. I found (Figure 17) that the Bayes factor favored the null hypothesis (that both sets were drawn from the same distribution) roughly 90% of the time. This shows that switch behavior cannot be driven solely by the frequency of long trials in the last few. In particular, it shows that the animal is not simply adjusting its switch point based on whether the last trial was long or short (this is the degenerate case, when memory size is 1). It is the amount that recent history *diverges* from more distant history that signals a change and is the key component to adjusting to changes quickly in the way that these subjects did.
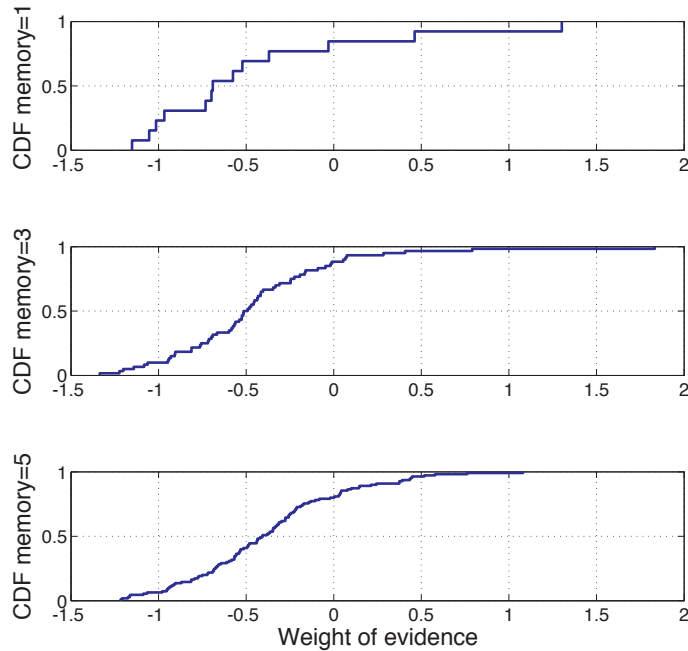
Figure 17. Cumulative distribution of the weight of the evidence in favor of the alternative hypothesis: that in comparing two groups of switch times (one with k long trials preceding it and one with j long trials preceding it where k ≠ j) the two groups were drawn from different distributions. The analysis is repeated with memory sizes 1, 3 and 5. The distributions clearly favor the null hypothesis.

If the switch behavior is not simply determined by the recent frequency of

long trials, this variable cannot be what is driving the shift in switch behavior after a

change in probability. Of course I do not claim that something beyond the sequence

of long and short trials (i.e. the stimulus) that drives the subject's behavior. Rather,

the simple models are missing a necessary feature: the ability to detect a change in

probability. This requires more than probability estimation; it requires storing a

probability value and comparing that value to a recent probability estimate–

determining the evidence that the current value is truly different from the historical

value.

Summary

I replicated the findings of Balci *et al.* (2009) in that the mice were near-optimal at the task and showed highly significant changes in their behavior in response to the change of the hidden variable. I expanded their analysis to quantify how long the change in behavior took after the change in the hidden variable.

Subjects detected changes in the underlying probability of a long trial soon after they occurred–which are what Dayan (2012) takes to be one of the hallmarks of model-based control. They then adjusted to that change rapidly, suggesting that subjects entertained distinct options for temporal decision criteria and moved to one or another when sufficient evidence had been accumulated, rather than gradually tuning their decision criterion to maximize their reward.

Behavior shifted quickly and abruptly, making it distinct from exploratory behavior. The changes were also made with very little negative feedback with which they might have judged one criterion against another. Their criteria were near optimal given the timing precision they displayed, despite the fact that the reward curves were often fairly flat at these precision levels. This means that they selected near-optimal criteria even when doing so did not have much effect on their earnings.

Finally, observed performance could not be simulated by simple models that did not take into account the subject's uncertainty or detect changes in probability. This pattern of results suggests a computational system that represents the relative probabilities of the two types of trials, evidence that a probability has changed and

the system's own uncertainty. These results are expanded upon in the experiments below.

Experiment 2: Behavioral shift latency is predicted by Kullback-Leibler divergence

Introduction

Experiment 1 showed that subjects shifted their behavior almost immediately in response to large changes in probability. This experiment expanded on that result by testing specific predictions about how the size of a probability change impacts the number of trials it takes for the subject to begin to shift their behavior. I will refer to this quantity as the shift latency. (Note the difference between switch latency and shift latency: the shift latency is the number of trials after a probability change that we see a shift in the distribution of switch latencies.) Naturally, we would expect larger probability changes to be responded to more quickly.

One principled way of measuring the difference between Bernoulli distributions (the probability of a short trial before the change vs. after) is the Kullback-Leibler divergence. The Kullback-Leibler divergence is a fundamental quantity in probability theory and information theory: it can be thought of as a kind of expected difference between two distributions (in our case, two Bernoulli distributions). More precisely, it is the information lost when one distribution is used to approximate another. If, for example, one thinks of a system optimized for distribution Q that now is made to operate over inputs from distribution P, the KL divergence is a measure of the inefficiency of that system (as opposed to one that

was actually optimized for P). In the case of our experiments, KL divergence can be viewed as the expected number of bits conveyed in each successive trial if you believe the probability of a short trial is p_1 and it is actually p_2. It is the rate at which evidence of a change is expected to accumulate.

Formally, the KL divergence from P to Q is defined to be

$$D_{KL}(P \parallel Q) := \sum_i P(i) \cdot \log_2 \frac{P(i)}{Q(i)}$$

The motivation for this measure and its use are obvious from this definition: for each possible signal (i) it is the information gained on $P(i)$ as opposed to $Q(i)$ (the binary log of the quotient of the densities at i) weighted by the actual probability of seeing signal i. In this way, even large differences between the two distributions have minimal impact if the differences are on inputs that are extremely unlikely. On the other hand, substantial differences between P and Q on a likely input will have substantial impacts on the KL divergence.

While the definition of this measure is intuitive, it has unintuitive consequences; for example asymmetry. If we take Q to be the true distribution as opposed to P, we end up with a different weighting of the difference of the logs. If Q is very different from P, this can lead to a very different KL divergence. Coming back again to our specific case, this means that in some cases the evidence that the probability value has changed will accumulate more quickly when the change is from p_1 to p_2 than vice versa. If subjects are using some criterion for the amount of evidence that must be accumulated that the probability value has changed before they shift their behavior, this should mean that we see shorter shift latencies for one change than for the opposite change.

Method

Because we generally only observe switches on long trials (subjects are usually fed on a short trial before they attempt to switch) and because the KL divergence between two Bernoulli distributions is symmetric around 0.5 $(D_{KL}(B(.2) \parallel B(.1)) = D_{KL}(B(.8) \parallel B(.9)))$ I decided to explore the half of the probability spectrum that would give us mostly long trials while still giving us a large asymmetry. This leaves us with a very simple question: do subjects show a shorter shift latency in response to the probability of a short trial being changed from 0.5 to 0.1 $(D_{KL}(B(.5) \parallel B(.1)) \approx 0.74)$ than in response to a change from 0.1 to 0.5 $(D_{KL}(B(.1) \parallel B(.5)) \approx 0.53)$? Conversely, we can also ask whether two changes with the same KL divergence would be reacted to in a similar timeframe, so compared the change with the smaller KL divergence above (0.1 to 0.5) to a change smaller in absolute magnitude but with approximately equal KL divergence: from 0.5 to 0.15 $(D_{KL}(B(.5) \parallel B(.15)) \approx 0.49)$.

Subjects were seven male C57BL/6j mice from Jackson Laboratories aged 6 weeks. As in Experiment 1, they underwent the same training and procedure described in the General Methods section: matching, autoshaping and switch. There were only two differences in procedure from Experiment 1. One, the long and short latencies were 4s and 12s during training and 4s and 8s during testing. Two, the proportion of short trials used in the switch sessions: alternating between 0.5 and 0.1 at first and then 0.5 and 0.15. Subjects were run at a given probability value until they had executed approximately 500 trials, at which point the probability value would change automatically, mid-session. This resulted in 5 +/- 1 alternations

between the probability values 0.5 and 0.1 and approximately 4 +/- 1 alternations
between the probability values 0.5 and 0.15.

Results

Just as in Experiment 1, for each probability change the data were fit before
and after the change. This time, however, we used a Weibull-Gaussian mixture
distribution rather than an exponential-Gaussian as in (Balci, 2007) because it
better fit some of our datasets. As before, the results were qualitatively the same
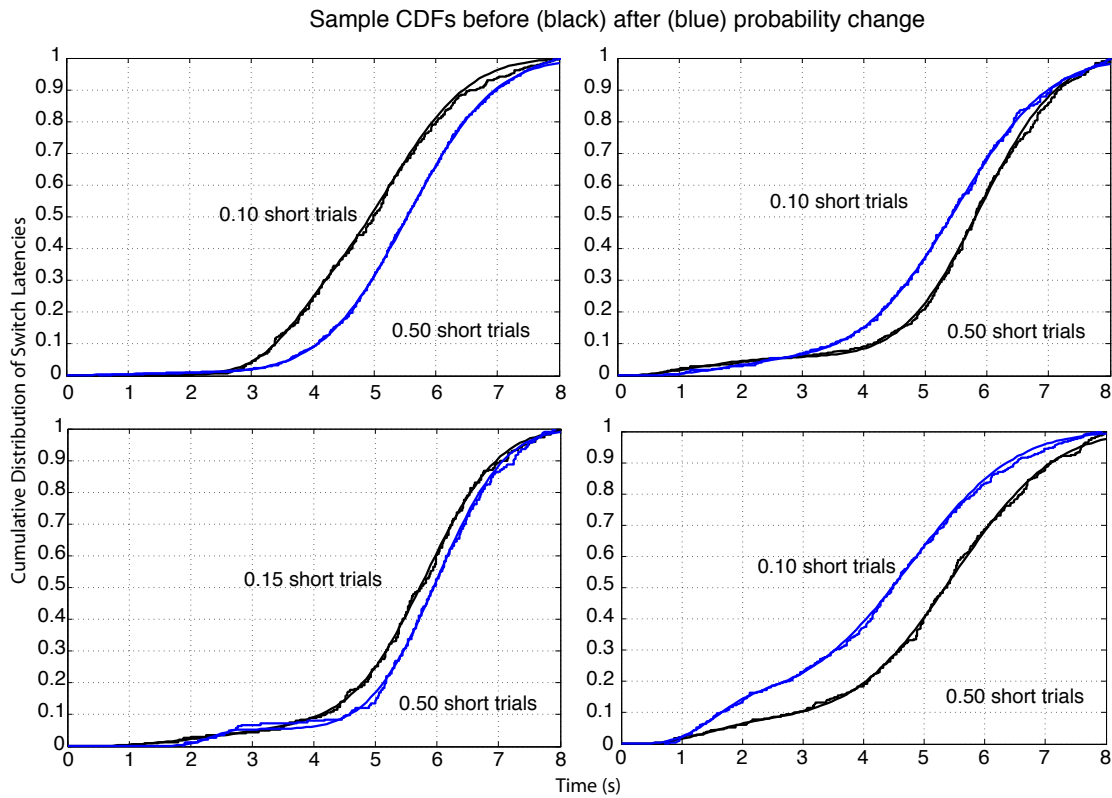using either fit. A few example fits can be seen in Figure 18.

*Figure 18. Four examples of switch behavior before (blue) and after (black) a change
in probability. The smooth curves are the Weibull-Gaussian fits and the jagged curves
are the raw data.*

Again as in Experiment 1, a two-parameter linear transition function was fit to the shift from the switch behavior before the change to the switch behavior after the change. Once again, the two parameters of this transition function were the shift latency and the number to trials to completion (gradualness) of the shift. I then examined the marginal likelihood function for the shift latency. Since we were interested in estimating the shift latency rather than choosing between two alternative models (as in Experiment 1) we took our summary statistic of the marginal distribution to be the expectation of the distribution. The summary of these shift latencies for all subjects can be seen in Figure 19. One can verify that the same pattern is also on the subject-level (and therefore not an artifact of averaging across subjects) by looking at the same data divided by subject in Figure 20.
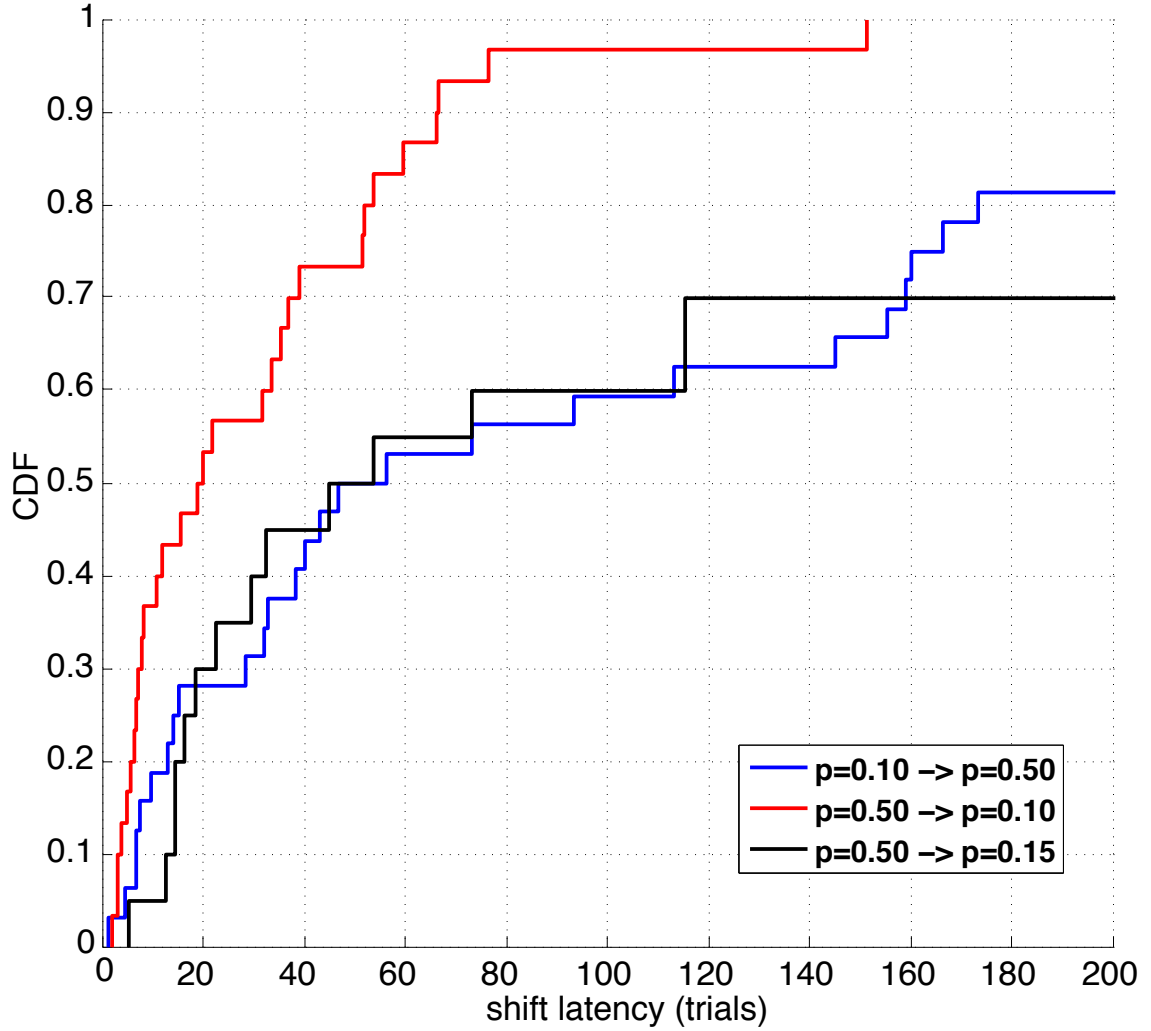
Figure 19. CDFs of the expectations on the shift latencies. As we would expect, transitions with a larger KL-divergence are responded to more quickly ($D_{KL}(B(.5) \parallel B(.1))$, in red). The other two transitions are not distinguishable but they also have approximately equal KL-divergences, so that is to be expected.
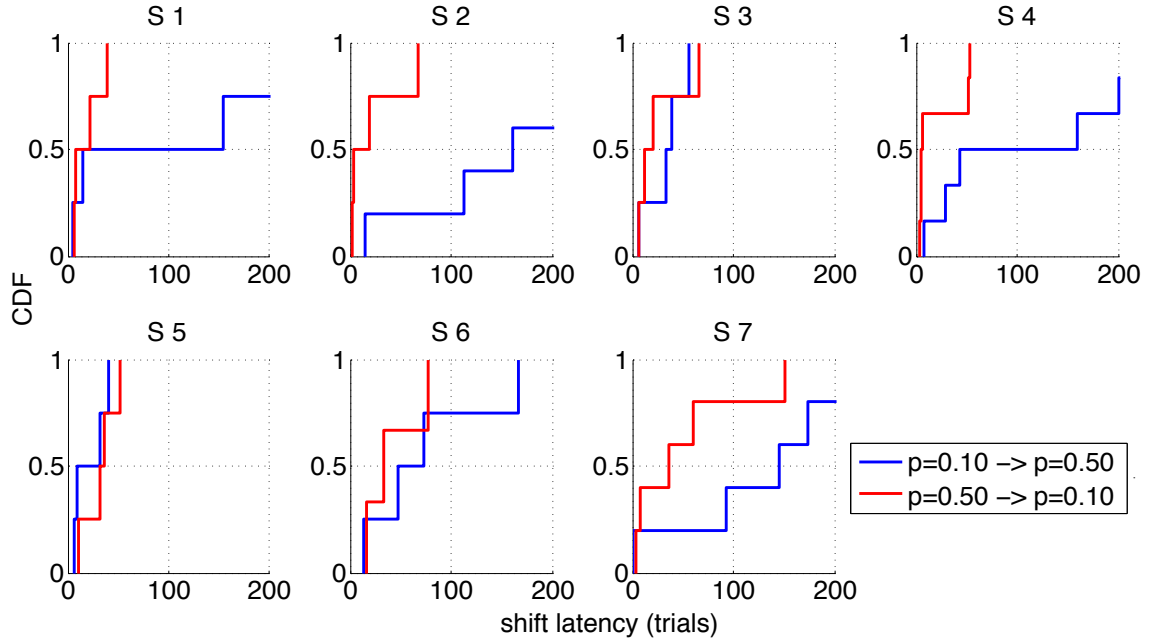
Figure 20. Expectations of the shift latencies divided by subject. Note that in all but one case, we see the same asymmetrical pattern as in the previous figure: subjects responded to a change from 0.5 to 0.1 more quickly than in the opposite direction.

The distribution for best estimates (expectations) for the shift latencies depends on DKL in the way one would expect: we see that the distribution changing from 0.50 to 0.10 has the shortest latencies, corresponding to the fact that $D_{KL}(B(.5) \parallel B(.1)) \approx 0.74$ is larger than $D_{KL}(B(.1) \parallel B(.5)) \approx 0.53$ because of the asymmetry in DKL. Finally, notice that the distribution when we move from 0.10 to 0.50 is intertwined with the distribution when we move from 0.50 to 0.15, which is predicted by the fact that $D_{KL}(B(.1) \parallel B(.5)) \approx 0.53$ is close to $D_{KL}(B(.5) \parallel B(.15)) \approx 0.49$.

It's worth noting that there wasn't a noticeable asymmetry with the values 0.5 and 0.15 despite the fact that $D_{KL}(B(.15) \parallel B(.5)) \approx 0.39$. One possibility for the lack of an asymmetry in the shift latencies when there was a difference in the KL divergence is simply that our method was not fine-grained enough to pick up on this small a distinction. The overlap is shown in Figure 21.
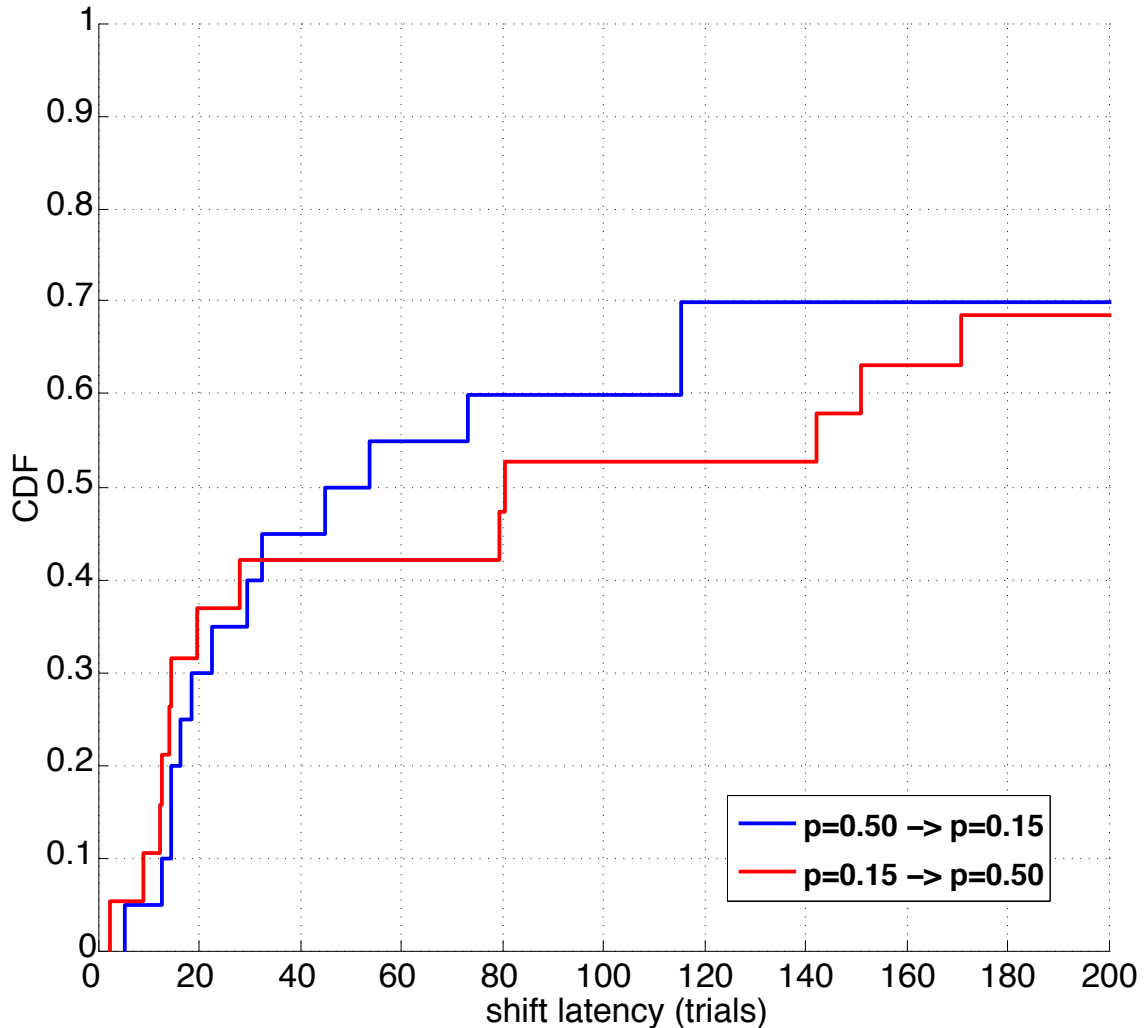
Figure 21. CDF of shift latency expectations for the smaller asymmetry: p=0.50 and p=0.15. Though there is an asymmetry in the KL-divergences of about 0.1, we do not see a noticeable change in the shift latency.

Discussion

We find that large differences in Kullback-Leibler divergence did predict

faster responses to a change in p. When framed in this way, the finding seems

blindingly obvious: Kullback-Leibler divergence is a measure of the expected rate of

accumulation of evidence for a change when a system was drawing from

distribution Q but is now drawing from P. If we make basic assumptions about the

criteria by which subjects decide they're dealing with a new distribution (see

General Discussion), it is almost self-evident that the larger the KL divergence of a change, the fewer trials it will take to pick up on that change. Those basic assumptions, however, are not made by most theories of animal learning.

Symmetry in other theories

It is not clear what assumptions are necessary to apply popular associative (e.g. Van Hamme & Wasserman, 1994) and temporal difference (e.g. Sutton, 1988) models of learning to the present task. These models operate by competition between the set of possible responses or predictions: the response that receives the highest level of activation from the environmental conditions is selected and executed. The learning with respect to each possible response is considered to be independent of one another. For example, an increase in the weight of the response "poke at the right hopper" does not affect the weight of the response "poke at the short hopper" (but of course the same experience can cause both of those weights to change). If we consider each possible target switch time as a possible response competing to be executed, it would seem that the weights of the possible responses would need to be related to one another. For example, learning that a target switch time of 6s is a bad choice should also impact the weighting of the target switch time of 5.9s. If instead we model the responses in competition as "poke at the short hopper" and "poke at the right hopper" and think of the associative weights as changing over the course of the trial, it is unclear how these weight are changing without reinforcement.

Rather than make large assumptions about how to apply these models to the current task, I will simply point out that their mechanism for parameter estimation

does not behave asymmetrically as our subjects do. As shown in Experiment 1, near-optimal performance on this task requires an estimate of the probability of a short vs. a long trial. Both of the models mentioned above use leaky integrators to estimate such quantities. At each trial, the estimate of a probability is the sum of a geometric series. The coefficients of each term is either one or zero depending on whether it was a long or short trial and the rate of decay is a constant determined experimentally. For example, if a subject had seen a sequence S of trials that was short-short-short-long-short, its estimate of the probability of a short trial would be[4]:

$$\hat{p}(S) = (1 - r) * \sum a_i * r^i = (1 - r) * (1 * r^4 + 1 * r^3 + 1 * r^2 + 0 * r^1 + 1 * r^0)$$

This leaky integrator varies about the probability value of the Bernoulli sequence it is input. Further, if the probability value is changed at some point in the sequence the old trials will continue to decay, eventually making a negligible contribution to the estimate of the probability value. The leaky integrator will therefore eventually vary about the new probability value. I will now show that the expected value of this leaky integrator does not depend on whether the sequence starts with p_1 and goes to p_2 or vice-versa.

Given a sequence of K trials drawn from p_1 $(a_1, a_2, \ldots a_K)$, followed by N trials drawn from p_2 $(b_1, b_2, \ldots b_N)$, the value of the leaky integrator is

$$\hat{p}(S_{1 \to 2}) = (1 - r) * \sum_{i=0}^{N-1} b_i * r^i + (1 - r) * \sum_{i=N}^{N+K-1} a_i * r^i$$

---

[4] This equation is sometimes presented without the normalizing constant $(1 - r)$ in front. I have normalized so that the weights sum to one here.

The difference from the true final probability is then

$$\hat{p}(S_{1\to2}) - p_2 = (1-r) * \sum_{i=0}^{N-1} b_i * r^i + (1-r) * \sum_{i=N}^{N+K-1} a_i * r^i - p_2$$

Since the terms $\frac{r^i}{1-r}$ sum to 1, we can multiply by this special case of 1

$$\hat{p}(S_{1\to2}) - p_2 = (1-r) \sum_{i=0}^{N-1} b_i r^i + (1-r) \sum_{i=N}^{N+K-1} a_i r^i - p_2(1-r) \sum_{i=0}^{N+K-1} r^i$$

$$= (1-r)\left(\sum_{i=0}^{N-1} b_i r^i - \sum_{i=0}^{N-1} p_2 r^i + \sum_{i=N}^{N+K-1} a_i r^i - \sum_{i=N}^{N+K-1} p_2 r^i \right)$$

The expected value of this difference (using the closed form of the sum of a

geometric series) is then

$$E(\hat{p}(S_{1\to2}) - p_2)$$

$$= (1-r) * \left(p_2 \frac{1-r^N}{1-r} - p_2 \frac{1-r^N}{1-r} + p_1 \frac{r^N - r^{N+K}}{1-r} - p_2 \frac{r^N - r^{N+K}}{1-r} \right)$$

$$= (p_1 - p_2)(r^N - r^{N+K})$$

Now if instead we reversed the probabilities so that it is instead K trials

drawn from p_2 $(b_1, b_2, ... b_N)$ followed by N trials drawn from p_1 $(a_1, a_2, ... a_K)$, the

value of the leaky integrator is

$$S_{2\to1} = (1-r) * \sum_{i=0}^{N-1} a_i r^i + (1-r) * \sum_{i=N}^{N+K-1} b_i r^i$$

and using the same simplifications as before, expected value and the true

probability is

$$E(\hat{p}(S_{2\to1}) - p_1) = (p_2 - p_1)(r^N - r^{N+K})$$

Which is just the additive inverse of the same quantity for the transition in the other

direction:

$$E(\hat{p}(S_{1\to2}) - p_2) = -E(\hat{p}(S_{2\to1}) - p_1)$$

This means that whether the transition is in one direction or the other, the leaky

integrator is expected to converge to the new probability value at the same rate.

Asymmetry and Bayes

Nearly any Bayesian model of animal behavior accords with our results here

because of the deep connection between Bayesian frameworks and information

theory, as explained in the General Discussion. Still, few researchers employing such

models have made predictions like the one in this experiment because their

hypotheses are often not framed in information-theoretic terms. To reinforce these

results, further experimentation should be done using KL divergence as a predictor

of how quickly animals will detect changes. Bernoulli distributions are relatively

simple univariate distributions but there is no reason not to apply the logic of this

experiment to more complex distributions, since KL-divergence makes no

assumptions about the two distributions being compared or even that they are the

same type of distribution. This experiment is and example of an information-

theoretic quantity being used to make a specific, unintuitive and yet correct

prediction about animal behavior. I take this as evidence that animal learning is best

understood in the lexicon of information theory.

Experiment 3: Squeezing temporal goalposts

Introduction

Experiment 1 established that subjects were proficient at detecting changes in probability and near ceiling at choosing a target switch latency that maximizes their rewards. This experiment explores the limits of this timing behavior by narrowing the short and long feed latencies (the "temporal goalposts" from Figure 2 of Experiment 1). Our immediate goal, then, is to establish an upper bound for the maximum timing precision of our subjects: we know they can discriminate 3s and 9s intervals with one level of variability, but do not yet know their variability when discriminating 4s and 6s intervals. At first, this may seem a technical question however, it informs larger theoretical issues.

There has been debate in the timing literature about whether there is a single timing mechanism or if different mechanisms work at different timescales (Gibbon, 1999; Lewis & Miall, 2009) and also about possible neural implementations of such systems–even though the behavioral properties of the system being described are not well understood (Buhusi & Meck, 2005). One of the cornerstones of this debate is whether subjects show a different coefficient of variation at different timescales. This is because the coefficient of variation in the output of a system is taken to be a physiologically meaningful property of that system (C. R. Gallistel *et al.*, 2013).

That is to say, the signature of variability could tell us a considerable amount about the calculation being implemented (Cordes, Gelman, Gallistel, & Whalen,

2001) and therefore what physical processes could be implementing it. On a larger

scale, insights into the signature of variability aid arguments as to whether or not

there is a single mechanism for representing all quantities, be they spatial, temporal

or numerical (Walsh, 2003). Finally, framing timing problems in an information-

theoretic context can aid us in deriving theories of the underlying computations, just

as there has been promising work framing working memory capacity in

information-theoretic terms (Brady, Konkle, & Alvarez, 2009).

I have thus far taken the coefficient of variation (standard deviation divided

by mean) to be a key measure of how precisely subjects can control their switch

behavior[5]. This is because a variety of experiments have showed the variability in

timing tasks seems to generally be proportional to the interval being timed, a

property known as scalar variability (C. Gallistel, King, & McDonald, 2004; Gibbon,

1977; Lewis & Miall, 2009). The fact that the coefficient of variation is constant is

often viewed as unsurprising as it can be thought of as the analog of a Weber

fraction in other areas of quantity estimation (Getty, 1975) although see (Lewis &

Miall, 2009)). This property can be seen in a variety of modalities and in a variety of

animals, even the yaw-gauge in houseflies (Rieke, Warland, De Ruyter van

Steveninck, & Bialek, 1997). In many cases, this property can be seen as the

consequence of sensory adaptation to keep within the dynamic range of the

---

[5] In keeping with the majority of the human timing literature, I here operate under the
assumption that for each subject, the CV does not vary with the duration of the interval being timed. In
fact, this kind of scalar variability (Buhusi & Meck, 2005; Gibbon, 1977) is the entire reason CV is discussed
in interval timing in the first place. After meta-analysis by Gibbon, Malapani, Dale, & Gallistel (1997)
suggested that CVs might differ for different ranges of intervals, this assumption has come into question. I
continue under this assumption, however, because there has been no consensus on even the direction of
the effect of interval length on CV and there does seem to be a consensus that the effect would not be
appreciable in the 4s to 12s range (Lewis & Miall, 2009). I will return to this assumption in the next
experiment.

communication pathway. This is analogous to scientific notation for numbers, which can represent a wide range of values with merely two digits, using one digit to represent the scale and the second to represent the measurement 0-9. Using this strategy, however, yields the same relative precision at all scales.

In this experiment, I squeezed the "temporal goalposts" (short and long feed latencies) to make the task more difficult and get a better upper bound on timing accuracy in this task. Subjects started on a very easily discernible pair of short and long intervals (4s vs 12s) and then experienced lower long feed latencies in subsequent sessions. I believe this experiment provides a new lower bound for the coefficient of variation for mice in interval timing tasks (one that is on par with typical human performance).

Because the primary goal of this experiment is to test the limits of how precisely subjects could time their switches, I focus my analyses on the Gaussian component of the Weibull-Gaussian distribution as the long latency decreases. As one would expect, the means lowered as the long latency decreased (Figure 23). Since the relative frequency of long and short trials was not manipulated in this experiment, the optimal target switch latency for a subject depended solely on the short and long latencies and the subject's timing accuracy. Ignoring the Weibull component of switch behavior, recall that the probability of missing a pellet is the probability of missing a pellet by switching too early when in a short trial plus the probability of switching too late when in a long trial:
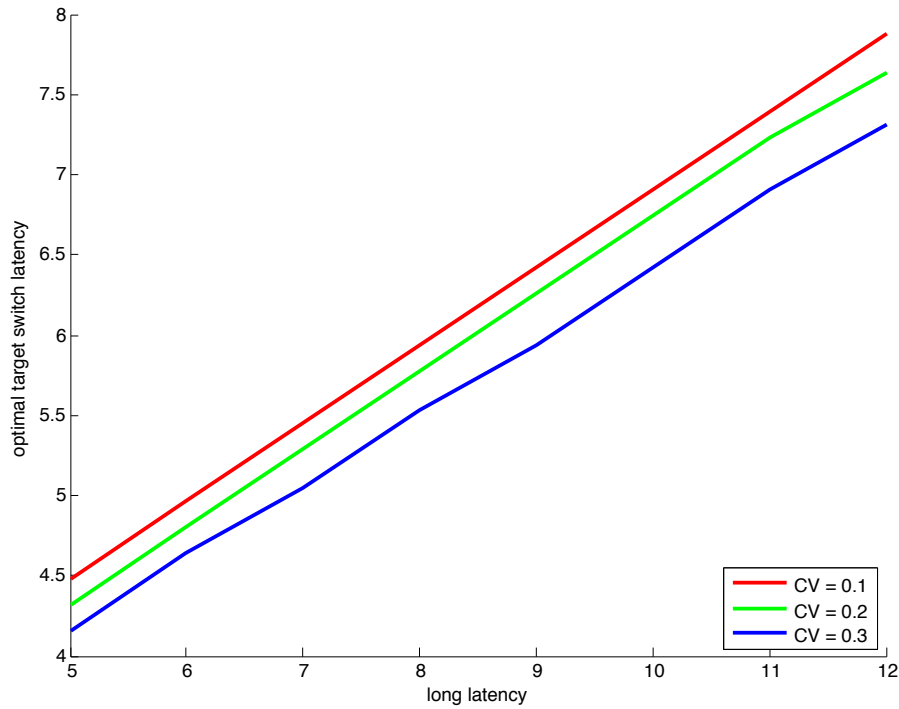
$$r(\mu) = \mathcal{N}(S|\mu, CV * \mu) * p_{short} + \left(1 - \mathcal{N}(L|\mu, CV * \mu)\right) * p_{long}.$$

Since to probability of short and long trials are fixed at ½ each in this experiment,

this equation can be simplified to:

$$r(\mu) = \mathcal{N}(S|\mu, CV * \mu) + \big(1 - \mathcal{N}(L|\mu, CV * \mu)\big).$$

The target switch time that minimizes the risk of losing a pellet for a given CV

decreases approximately linearly as the long latency decreases (Figure 22).

Figure 22. Optimal target switch time as a function of the long latency (short latency is fixed at 4 seconds). Note that for any of the plausible CVs, the curves are approximately linear: the lowering of the longest acceptable switch time means that subjects must aim to switch earlier.

Method

Subjects for this experiment were 21 female L1-knockout mice (L1CAMKi, 10

homozygous and 11 heterozygous) obtained from Schachner University, Hamburg.

These mice were originally obtained to examine possible deficits caused by this

knockout but their performance in relevant measures did not exclude them from

study here. In fact, counter to what one would expect, the heterozygous mice were

better timers in many cases than wild type littermates. The genetics of this are not the focus of this experiment and all relevant effects were seen in both groups, so I have collapsed over type in the analyses below.

Subjects experienced the same matching, autoshaping and switch procedures described in Experiment 1 except for the feed latencies: for the duration of the experiment, the short feed latency was fixed at 4 seconds while the long feed latency was initially 12 seconds and then varied as our experimental manipulation in subsequent sessions. The long feed latency values were 12s, 8s, 7s, and 6s.

Results

Unsurprisingly, we indeed see the Gaussian means steadily decreased as we decreased the long latency, as seen in Figure 23 (though the means were higher than the optimal target switch latencies in Figure 22 above). Note that a decrease in the mean requires an even larger decrease in standard deviation in order to lower the CV. So, when we see a lowering of the CVs (Figure 24) as we decrease the long latency, this represents not only a tightening of the standard deviation, but a tightening disproportionate to the lowering of the mean.
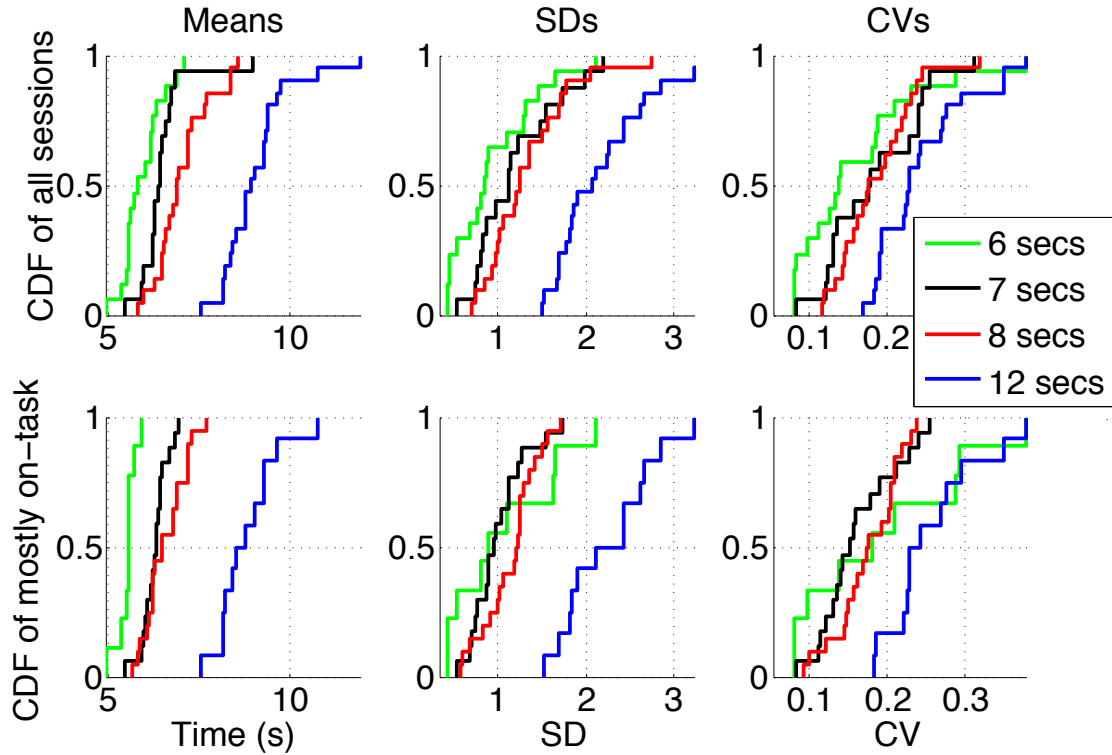
Figure 23. CDFs of the means and standard deviations of the Gaussian components of the switch behavior and the corresponding coefficients of variation (SD/mean). These are displayed both before and after throwing out sessions in which the subject was not on-task at least 50% of trails. The observed mean switch times decreased as the long latency decreased, in accordance with the optimal behavior outlined above. Subjects seem to be able to decrease their CV when required to do so (though the 8s and 7s cases do not display very different CVs, the effect is in the correct direction and all other comparisons are clear), but performance breaks down (for SD but not mean) in the 6s condition.

We now come to the main result of this experiment: CVs decreased as long latency decreased (Figure 23, right panels). Before we may conclude that this represents the subjects actually timing more precisely, we must rule out a few alternative explanations. One we have already ruled out is that the decrease in CV is caused by the mean increasing: we have seen that the mean is in fact decreasing. Again, if the CV is decreasing despite the fact that the mean is also decreasing, the standard deviation must be decreasing more rapidly than the mean is decreasing.

Another possible explanation for our results is a practice effect, which can improve the estimation of temporal intervals, as with musicians (Grondin & Killeen, 2009). While it is true that I decreased the long latencies sequentially, meaning that subjects were more experienced when they reached sessions with lower long latencies, the CVs here are lower than in the previous experiments (even ones in which subjects experienced far more trials such as Experiment 1). So, it seems practice alone cannot explain the decrease in CVs we see in these data. I also checked this by looking for differences between the first half of trials within a session as compared to the second half from the same session. I failed to find any reliable difference there.

It therefore merely remains to check that the decrease in standard deviation is not due to some artifact of the fitting algorithm used. The primary candidate for this type of artifact would be a situation in which the fitting algorithm selectively narrowed the switch latencies it ascribed to the Gaussian component of the mixture distribution. Recall that some switch latencies are ascribed to the Weibull portion of the distribution and these are taken to be "not on-task" because they are switches in which the subject almost immediately switches to the long hopper, apparently without timing an interval. Similarly, some proportion of the long trials that the subject misses by waiting too long at the short hopper are also "not on-task" for the opposite reason: the subject has started grooming, went back into the home cage, etc. So, the primary way in which the decrease in CV could be a product of the fitting algorithm would be if the algorithm classified too many trials as not on-task *and*

chose abnormally steep portions of the distribution to estimate the standard

deviation of the Gaussian distribution.

A natural way to limit this possibility would be to throw out all sessions in

which the fitting algorithm ascribed a high proportion of the trials to be not on-task.

Undoubtedly, this will also exclude sessions that have been fit accurately, but this

technique should be sufficient for assuaging worries that the observed effect is due

to fitting artifacts. Unfortunately, this also eliminates the majority of sessions in the

6s condition because subjects timed-out on a large portion of trials in these sessions.

While it is understandable that timing performance would break down with such

narrow temporal goalposts (like making the goal too small in soccer), this means

that a large portion of the distribution of switch times is extrapolated (e.g. Figure

24).
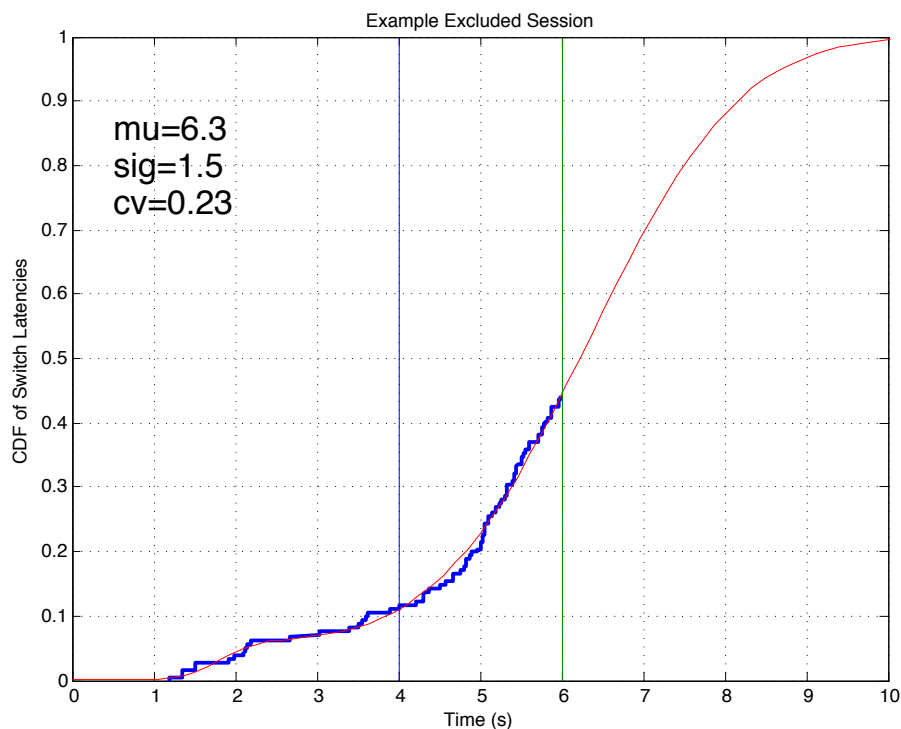
Figure 24. Example session excluded because of too many time-out trials. In this session, the subject timed out over 55% of long trials (stayed at the short-side hopper for six seconds). The shape of the uncensored distribution allows us to extrapolate the rest of the fit function but basing conclusions on so much extrapolation can be difficult.

Discussion

The most theoretically important result is clear: subjects display a different coefficient of variability in their timing when their task is made more demanding. Thus, a subject's CV isn't fixed, even in this limited temporal range, as has traditionally been assumed (Gibbon, 1977). The details of our interpretation, however, depend somewhat on whether or not we include sessions in which the subject timed out a large portion of trials.

If we do not exclude such sessions, the afore-mentioned result can be seen in all four conditions of our experiment: the CVs from the 12s, 8s, 7s and 6s long latency conditions progressively lower. This means that timing is not subject to

Weber's law which asserts that a given subject's accuracy (as a proportion of the quantity to be estimated) is constant. The dramatic increase in the number of time-out trials in this condition, however, suggests that the 6s long latency condition may be the limit of how far this task may be pushed to yield lower CVs.

On the other hand, if we do exclude sessions in which half or more of the long trials were timed-out, we face a difficult problem. Any reasonable criterion for the proportion of time-outs tolerated eliminates a large number of sessions. Even the lax criterion of requiring that half the long trials not be time-outs excludes nearly half of the sessions in the 12s condition and the 6s condition (though virtually none in the 8s and 7s conditions). This isn't necessarily a problem in itself. For example, in the 12s condition the elimination of these trials does not change the shape of the distribution of means or standard deviations (the blue lines in plots 3 and 6 of Figure 23 are similar despite the fact that one is comprised of half as much data). We interpret this to mean that the excluded trials were simply ones in which the subject hadn't yet grasped the task: the thrown-out sessions were representative in their distribution of parameters.

In the 6s condition, however, the trials excluded have a large impact on the shape and location of the distribution of standard deviations (though not of means): the thrown-out sessions were largely ones with low standard deviations. This means that after these sessions are excluded, the SD and CV distributions for the 6s condition overlap heavily with the other conditions. The most obvious interpretation of these data is that excluding the sessions with too many time-outs

has revealed the true behavior in the 6s condition: performance simply breaks down and one sees a high variability in CVs.

The problem with this interpretation of the data is that the means in the 6s condition were not similarly impaired. The exclusion of the sessions in question did not lead to a change in the shape of the distribution of means, which is tightly centered around ~5.5s. If we believe that timing behavior in the 6s condition is indeed more variable than in, say, the 7s condition, we must accept that this is an instance in which the standard deviation of time estimates becomes decoupled from the interval being estimated.

Under either interpretation of switch behavior in the 6s condition, it seems clear that subjects are capable of influencing the coefficient of variation in their timing when the task requires it. Further study is required to see if this result generalizes to other modalities of quantity estimation and what task parameters motivate subjects to influence their CV. The truly remarkable result here is that one can find a reliable change in the CV of a single subject, not at "break points" in the timing mechanism (where one physical process takes over for another, better suited to a different scale as (Gibbon *et al.*, 1997) suggested might occur at 0.1s and 1.5s) but rather in order to adapt to the task before them.

Experiment 4: Variable feed latency

Intro

A key aspect of the analyses Experiments 1 and 3 was the amount of variability in the switch latencies. I see this as analogous to results in motor planning that similarly show humans are sensitive to their own motor variability when planning where to point (J. Trommershäuser *et al.*, 2008). The studies on motor planning have been careful to differentiate between various sources of variability. For example, when quickly pointing at a target on a touch screen, there is some variation in the locations actually touched because of variability endogenous to the motor system. This is also variation in the point recorded by the touch screen when a given point is actually touched. There is also variation in the point recorded by the touch screen when a given point is actually touched. This is exogenous to the motor system but is nevertheless relevant in motor planning when the subject's performance is based on the endpoint recorded. Trommershäuser, Gepshtein, Maloney, Landy, & Banks (2005) artificially increased this exogenous variability and found that subjects learned to account for the artificially inflated total variability so as to reach near-optimal pointing targets. The design of the current experiment similarly added exogenous variability to the system being studied but yielded different results.

The adjustment subjects made to their timing behavior did not yield optimal responses for the new, combined level of timing variability. They did, however,

respond to the increase in exogenous variability in a way that suggests that they are sensitive to the variability in their timing mechanism itself, not merely the total variability in the durations apparently experienced. I will argue that this implies that subjects distinguish between variability in their measurements and variability in the durations they are measuring.

Note that just as with the end points recorded by a touch screen, the set of duration measurements that a subject stores in memory has two distinct sources of variability: endogenous variability, which is the variability of the timing systems we have discussed at length thus far, and exogenous variability, which is the variability in the durations themselves. Because our long and short feed latencies have been constant thus far (at least within each session) this issue has not yet come up: the short feed latencies have been (for example) 4s with negligible variability, so the variability in the recorded interval measurements has been identical to the variability of the timing system. This experiment added variability to the intervals being timed, thereby increasing the variability in the interval measurements recorded by the subject, and explored how this exogenous variability changes subject behavior. Specifically, I argue for the somewhat surprising result that models of timing must represent these two sources of variability separately rather than simply keeping track of a combination of the two (recorded duration measurements).

The most common models of timing (e.g. Gibbon, 1977; Meck, 2003) make no distinction between exogenous and endogenous sources of variability. I will make a case study of the failure of Scalar Expectancy Theory (SET) (Gibbon, 1977) to

explain the results of this experiment because it is still the most influential model of timing. My comments, however, will be easily applicable to the broad swath of models that do not distinguish between endogenous and exogenous variability. These include popular behavioral models such as The Behavioral Theory of Timing (Killeen & Fetterman, 1988), Learning to Time (Machado, 1997) and memory-trace models such as the Multiple Time Scales model of Staddon and Higga (1999).

The problem

First, I will set lay out how difficult dealing with timing variability is and how SET treats it. Time is a unique and inherently tricky domain as has been noted by physicists and philosophers alike (for a nice review, see Montemayor, 2012). For one thing, timing stimuli can only be presented once. While one can examine the same object from different angles and in different contexts (though admittedly, never again in *exactly* the same way), a duration can only be experienced once. Since variability is necessarily a property of a set of observations, determining the variability of a class of durations can be difficult as opposed to determining the variability in the lengths of a set of objects. One can have a shoemaker generate a hundred shoes of the same size and compare their lengths to get a sense of that shoemaker's variability but one cannot perform such a side-by-side comparison of the amount of time taken to make each shoe directly to one another.

Estimating the variability of one's own timing system is a particularly tricky instance of this problem: one can measure the length of an object multiple times to determine the variability of the measurement process but there is no clear analogue in the temporal domain. Since direct comparison of asynchronous intervals is not

possible, indirect comparison is necessary to establish the temporal variability of a process. This is not problematic for humans because we have objective duration-measuring processes so reliable[6] that we can consider time to be a uniform and continuous objective quantity at the scale we are considering.

That is to say, we do not need to worry about our standards of time: even a cheap quartz watch provides such accurate and reliable timing that its variations are completely negligible on our timescale, so give us a temporal gold standard. We know that seven ticks of the second hand of our watches are equivalent, so we can compare asynchronous intervals by noting how many ticks each of them took. Mice, however, do not own quartz watches.

If direct comparison of asynchronous durations is not possible, calculating the variability of a set of them requires the assumption that time behaves the same way during one duration as during another. This requires a concept of objective time: it is essentially using an allocentric time coordinate system. I will argue that mice distinguish between endogenous and exogenous timing variability and that they therefore implicitly encode times as an allocentric temporal coordinate. In a very real sense, then, the decision-making system I am studying operates under the assumption that time is independent of their perception.

It is easy to see why the vast majority of theories of timing would not posit such a sophisticated concept of objective time. Mice do not have meetings to get to at 3pm, so it is not obvious why they would need to know the variability of their

---

[6] It's rather tricky to rigorously define what makes a reliable and accurate temporal standard but I here take it to mean something akin to "gives approximately the same result repeatedly when used to measure processes that also are considered to be good temporal standards."

timing. Why they would calculate their temporal variability is not immediately obvious given the simple day-to-day foraging we imagine is the primary problem mice face. For a wide variety of tasks, one could make do with an egocentric temporal map with no need for an allocentric temporal representation. Further, it's not obvious how they would keep track of objective time even if doing so would be advantageous. They do not have timekeepers that keep objective time in nature; even the amount of daylight in a 24-hour period varies greatly over the course of a year. The lack of temporal gold standards, like the ticking of a quartz watch, in nature poses a puzzle about how mice could keep track of objective time.

Popular treatments

Most models of animal timing avoid this problem by simply not representing endogenous variability. Take as a primary example, SET. SET models timing behavior using a pacemaker-accumulator, a memory and a comparator. The pacemaker is a sort of metronome, much like the pulsing in a cesium beam clock. The amount of time it takes to pulse is usually assumed to be short and Poisson-distributed with relatively low variability deviation. Because the accumulated pulses of such a process would be binomial-distributed, which does not accord with Weber's law, the pacemaker is further assumed to run at a variable rate from trial-to-trial, which ensures scalar variability. At the start of each trial, the pacemaker-accumulator is zeroed and pulses begin to accrue for the duration of the trial, until the subject is reinforced. The final reading from the accumulator is then recorded to memory. In future trials, the decision of whether or not to respond at a given time is made by comparing the current reading of the accumulator against a sample from

memory of how many pulses were accrued before reinforcement. Typically, this decision process is simple, as in "respond if the ratio of the current reading to the sampled reading is greater than b" for some constant b. Only the perceived duration is stored with no further calculation.

The variability in the subjects timing behavior comes from the fact that the decision criterion is a comparison of the current perceived duration against a sample from memory. Since there is variance in the stored perceived durations, this (plus the endogenous variability acting on the current duration) causes the variability in the subject's timing behavior. When we add variability to the intervals being timed we will be adding variability to the perceived durations and therefore should, according to SET, see increased variability in timing behavior.

Other popular models of animal timing also make no distinction between exogenous and endogenous variability because they are (just as in SET) not explicitly represented. The Behavioral Timing (BeT) model operates by an internal pacemaker switching the subject from one behavioral state to another. When setting up this model, Killeen and Fetterman (1988) explicitly point out that their model makes no reference to the subjective perception of time or internal scales of it. Because of this, it obviously cannot represent endogenous variability in the timing of intervals and therefore cannot be sensitive to it. More specifically, the addition of variability to the time of reinforcement would increase the probability of being reinforced when in various different states. That would mean more states becoming associated with reinforcement and therefore greater variability in timing behavior. This is exactly the opposite of what our data show in many cases below. I will now

touch on how the Learning to Time and Multiple Time Scales models similarly

predict greater timing variability because they do not explicitly represent

endogenous variability.

The Learning to Time model is similar in that activation spreads serially

from one "active state" to the next without the subject explicitly representing time

or recording intervals. When a trial begins, the first state in the subject's chain of

states becomes active and as time passes, activation goes from one state to the next.

The intermediate states may or may not trigger specific behaviors, while the

terminal state does trigger the operant behavior. The strength of the connection

between a state and a behavior changes as in typical models of learning by

association: chains of states that earn reinforcement are strengthened while those

that fail to earn reinforcement are weakened. This model is also would increase

variability in its output when timing variable intervals.

For memory trace models, such as the Multiple-Time-Scale (MTS) model,

timing behavior is explained as the comparison of memory traces, similar to the

leaky integrator described in Experiment 2. In this model, whenever an event is

experienced, a memory is created with a given strength. As time progresses, the

memory weakens by some proportion at each instant. Provided that one can recover

the initial strength of the memory (the simplest case would be in which all

memories initially have the same strength and this is hard-coded), animals could

then mimic timing behavior without actually having an internal timer by associating

a behavioral response with a below-threshold memory trace of the stimulus onset.

For example, if a feeder is armed two minutes after a signal light coming on, the

subject would learn to associate poking in the hopper when it has a weak memory trace for the light coming on (weakened by two minutes of decay) with a reinforcement. Then, when the reinforcer is delivered, the strong trace of the "reinforcer delivered" memory would become associated with not poking in the hopper because such attempts would go unreinforced until the signal light had come on again for two minutes. Thus, memory strength would cause the subject to cycle through various behaviors associated with different memory strengths, making the subject appear to be timing the interval without use of an internal clock.

Once again, this theory predicts that increasing feed latency variability would also increase timing variability. If a memory trace is not strengthened regularly, but rather only approximately regularly, the time at which the trace of being reinforced drops below threshold will also become more variable. This would then mean that we should expect to see increased variability in the timing behavior of subjects.

I find a number of interesting and unexpected results; the one of most obvious theoretical import is that in many cases increasing the variability in the intervals lowers subject timing variability. I will argue that this means that either subjects are using some other quantity as a gold standard or they have a built-in method for assessing their timing variability independent of stored perceived trial times.

Method

The setup of this experiment is largely similar to Experiment 3 above except rather than manipulating the value of the long feed latency, we used variable long and short feed latencies. Subjects were first trained on the standard matching,

autoshaping and switch paradigms with feed latencies 3s and 9s. After this, they were introduced to Gaussian-distributed short and long feeding latencies. Subjects who were still struggling with the task were given latencies with mean still 3s and 9s respectively but whose CVs were manipulated. Once subjects were proficient at the task, they experienced feed latencies with means at 4s and 8s whose CVs were manipulated.

Since timing behavior is generally Gaussian-distributed, we also made the feed latencies Gaussian-distributed. If, as SET would predict, timing behavior is Gaussian-distributed because the stored perceived durations are Gaussian-distributed, adding Gaussian variability to the durations being timed should be computationally indistinguishable from the animal's endogenous variability increasing by the same amount. The distribution of perceived durations should look like the animal was having particular difficulty timing that day. This is because the noise distribution was a Gaussian with mean zero and the sum of two Gaussians when one has mean zero is a Gaussian with the same non-zero mean and variance that is the sum of the two variances. If the distributions of the exogenous and endogenous variability were of different forms, it might be possible to separate the two after the fact by methods similar to those we have used to separate the impulsive switch behavior from the timing behavior. Of course, SET does not predict any such parsing of the perceived durations.

Subjects were 29 male mice of species C57BL/6J. They were approximately 7 weeks old when the experiment began. Of those, 5 were excluded from analysis because they never achieved proficiency on the training tasks.

The probability of a long trial was again fixed at 50% so as not to bias animals to one side or the other. We manipulated two variables in this experiment: the location of the Gaussian-distributed feed latencies (short, long or both sides) and the CV of the Gaussian distribution used (0.1, 0.2 or 0.35). Thus, there were 9 possible conditions: three levels of CV in three possible locations (Figure 25). With even a modest number of sessions, the number of possible orders of conditions grows quickly, so we could not explore the effect of order exhaustively. Instead, we chose a few paradigmatic orders to explore. They were: increasing the CV (a single subject might see 0.1L, 0.2L, 0.35L), decreasing (e.g. 0.35R, .2R, .1R), or manipulating only the location (e.g. .1L, .1R, .1B). No significant effect of the order chosen was found so all reported analyses were performed collapsing across order.
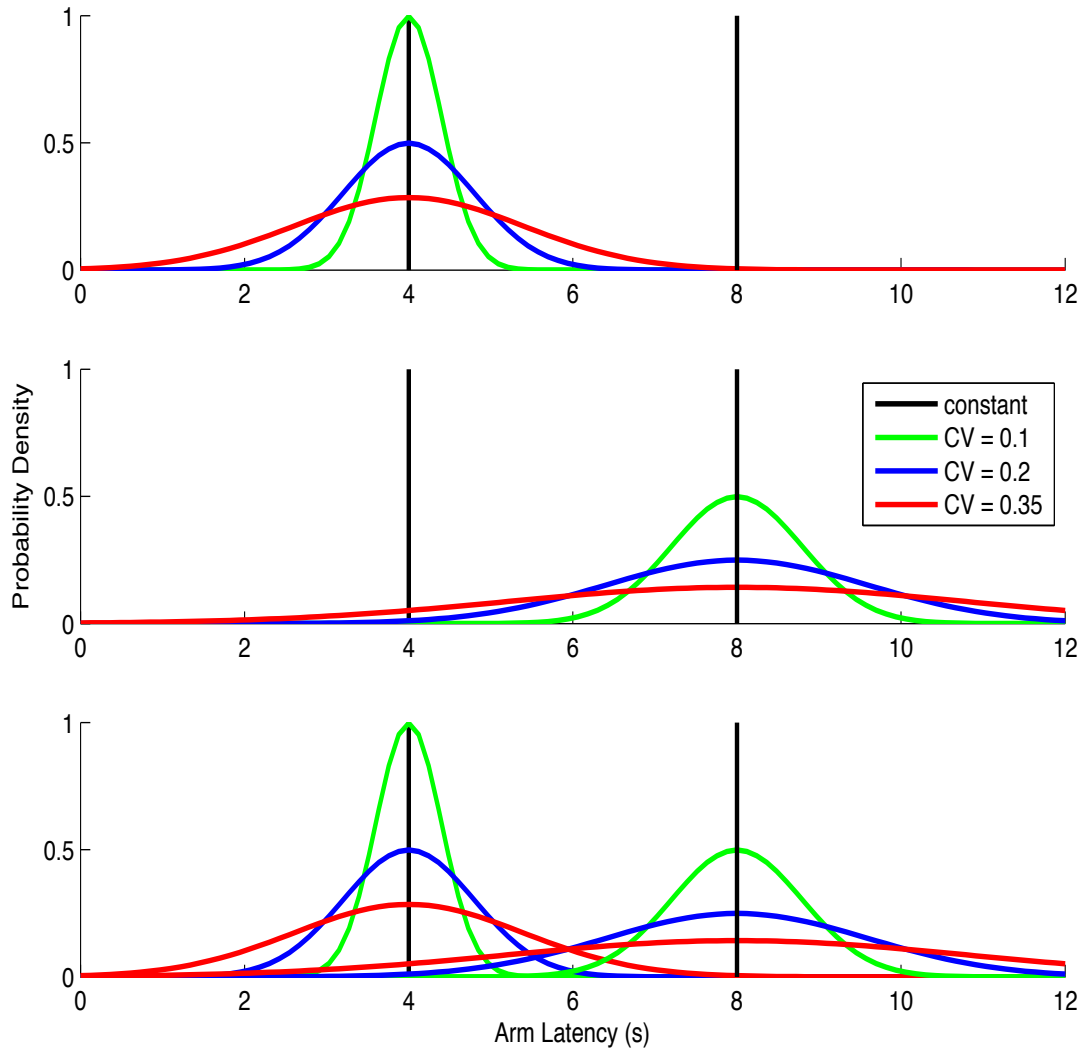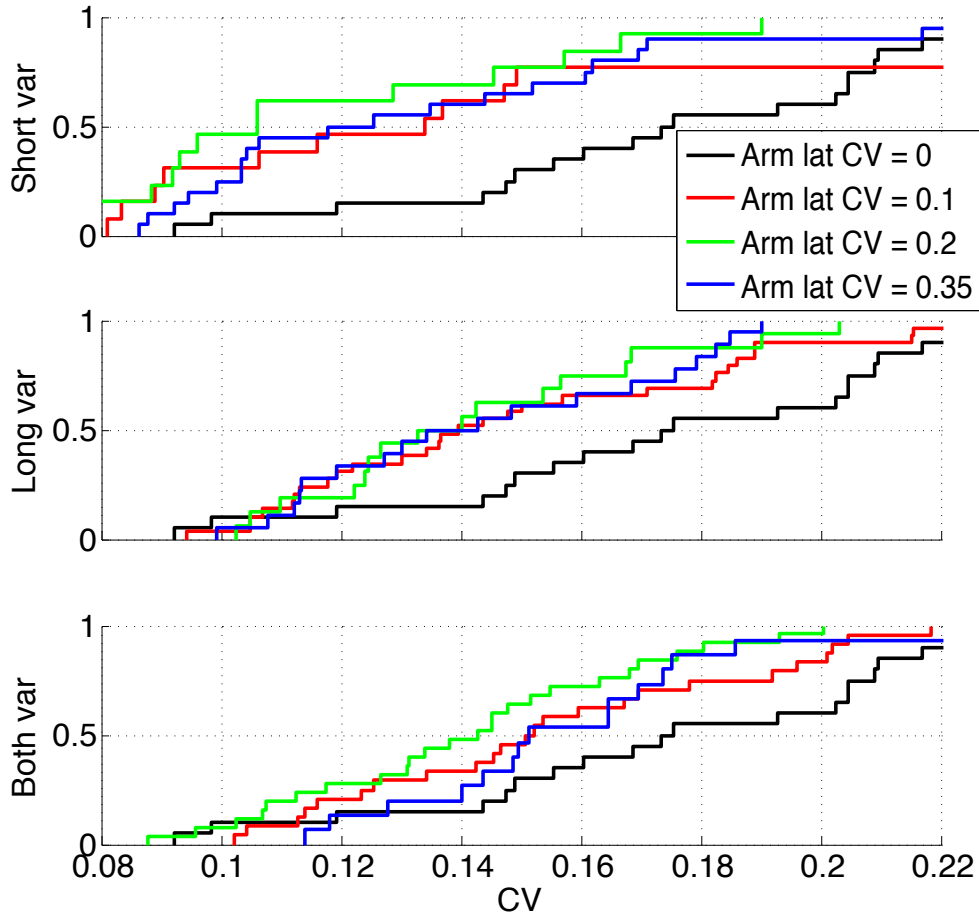
Figure 25. Distributions of switch latencies in each condition. Note the difficulty of the task in the CV = 0.35 condition: there is a huge overlap between the distribution of short and long arm latencies. In this case, it wasn't uncommon to have long arm latencies that were shorter than some of the longer short arm latencies.

Results

Two things are immediately obvious by inspection of the CDFs of the switch latency CVs (Figure 26): First, the CVs are noticeably lower than in the constant 4s vs. 8s condition (taken from Experiment 3). In every case, the conditions with variable arm latencies have a distribution of CVs that are to the left of the constant arm latency condition. This is especially true when variability was only added to the

short arm latency. In addition, increasing arm latency variability from 0.1 to 0.2 seems to always decrease timing CVs, though there does not seem to be further improvement when moving from the 0.2 to the 0.35 level. Second, for each variability location, the level of the variation does not seem to have a large impact on the timing accuracy.


Figure 26. Cumulative distributions (across subjects) of the coefficients of variation in the distributions of switch latencies as functions of level of variation in the feed latencies (colors within plots) and which latency varied (between plots). Random variation in the objective durations reduces the coefficient of variation in timed switches rather than increasing it as would be expected on any model that, like SET, does not distinguish objective variation in duration from subjective variation. This result holds regardless of the location of the variation (short-only, long-only or both sides) and the level of variation.

The result of primary theoretical import is that adding variability to the arm latencies lowers the variability in the timing of the switches. I will argue that this means that subjects are separately representing endogenous and exogenous sources of variability. This will be the primary result of this experiment and I believe it demonstrates the level of representation subjects use to make decisions under uncertainty.

There are, however, a number of surprising results in the details of our data that merit exploration. First, it is surprising that the switch latency CVs are higher when the feed latency CVs are 0.35 than when they are 0.2. As we saw in Experiment 3, making the timing task more difficult by squeezing the temporal goalposts resulted in subjects lowering their switch latency CVs. In the current experiment, the standard switch task is also made more difficult: a given level of timing precision in the constant 4s vs 8s condition would lead to more missed trials in the 4s vs 8s variable condition. The higher the variability in the arm latencies, the more expected missed trials. In the range of values most of our subjects fall into, the same switch behavior in the constant feed latency condition would lead to 2-3 times as many missed trials in our highest variability condition (feed latencies vary by 0.35 of their mean). We see that switch latency CVs decrease when the feed latency CVs increase from 0 to 0.1 and again from 0.1 to 0.2 but not from 0.2 to 0.35. This may be attributable to a breakdown in performance because of the difficulty of this condition.

Second, the means also do not respond as expected: the means of the timing behavior does not track the optimal mean for the increased level of variability. We

calculated the optimal mean for each session to be the mean that would maximize

the number of reinforced trials given the subjects' CV for that session. If we restrict

our attention to the "both" location, we see from a scatter of the optimal means

against the observed means (Figure 27, left panels) that as the variability of the feed

latencies increases, the observed means drift away from the optimal means. As we

can see from the CDFs of the same data (Figure 27, right panels), the spread of the

observed means does not increase. When the variability of the feed latencies is

increased, the means of the switch latencies creep upward even as the optimal

means decrease.
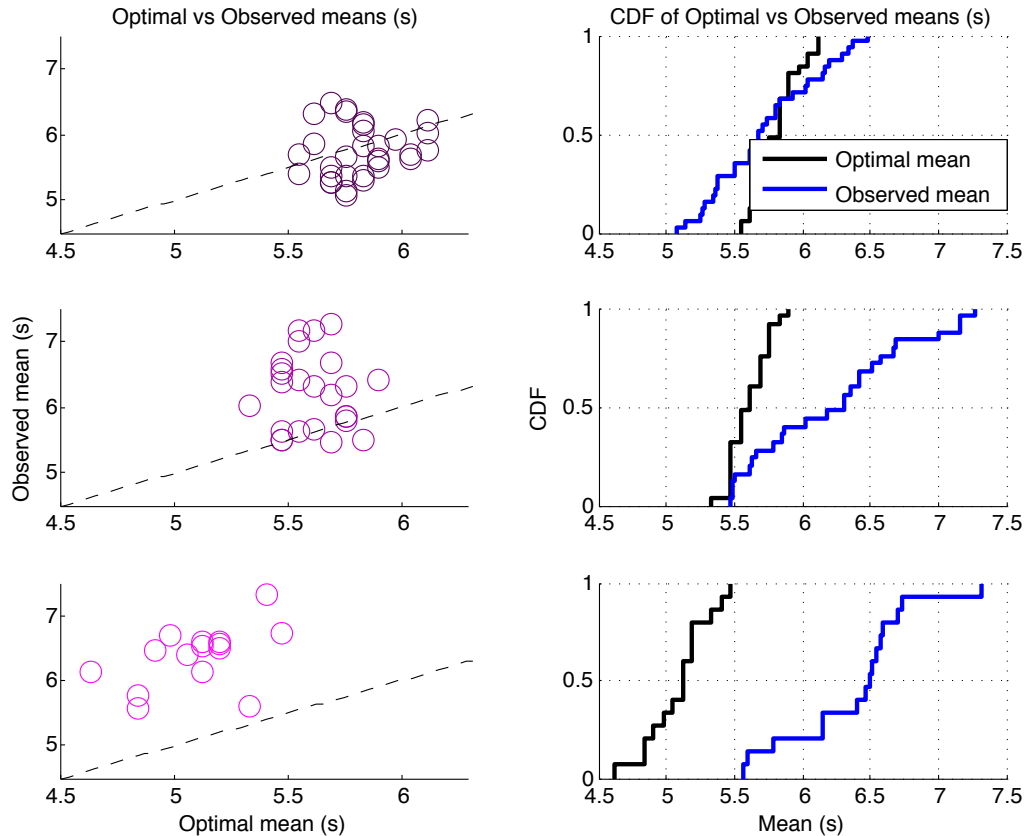
Figure 27. Left column. Scatter plot of observed mean switch latencies versus optimal mean switch latencies and the identity line (dashed black) around which the data would cluster if the means of the switch latencies were approximately optimal. The level of jitter in the objective feed latencies increases from top plot to bottom plot. Right column. Cumulative distributions of observed mean switch latencies (blue) and optimal mean switch latencies. Notice that the blue plots shift rightward as one goes from top to bottom, whereas the black shift leftward. All sessions shown here varied the feed latency at both hoppers and sessions are grouped by the level of variation (top: CV = 0.1 middle: CV = 0.2 bottom: CV = 0.35).

So, plainly subjects are not simply optimizing their behavior for the new, combined level of variability. They also aren't systematically lowering their CVs as the task becomes more difficult in some locations (higher variability conditions). When we look at the timing latency CVs in comparison to their means (Figure 28), we find that a) when only the long feed latency varies (middle row) we do see basically flat CVs (scalar variability), but when the short or both feed latencies vary, we see a negative correlation between mean and CV, and b) it appears to be

approximately the same correlation in all those cases. If this is a reliable effect, we

are seeing decreasing relative variability, rather than scalar variability. Of course,

this trend cannot continue indefinitely because we are reaching the bottom of

plausible CVs for our subjects. Thus an exponential fit is theoretically more plausible

and accounts for more of the variance than a linear one (Figure 29). I cannot think of

a plausible model that predicts the pattern we see here, in which the CVs are

negatively correlated with the means.
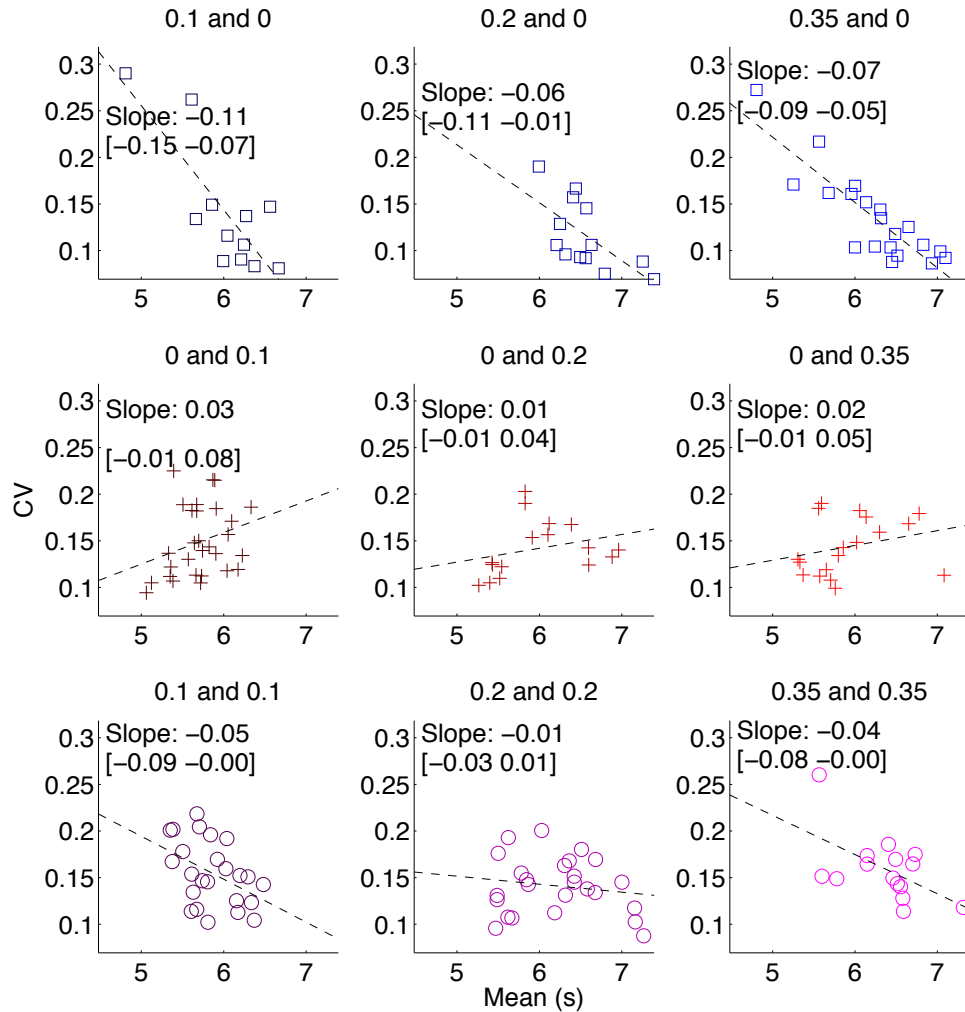
Figure 28. Mean-CV pairings sorted by condition with regression lines and confidence intervals. Again, we see that the level of variation does not matter as much as the location of the variation. In addition, we see that when the left side or both sides are variable, there is a negative correlation between the mean and the CV of the switch behavior.

So, while the long-only conditions show a weak positive correlation between CV and mean (all three just under 0.3), the other six conditions display negative correlations. Taking these six as a group, the trend remains strong (Figure 29) with a negative correlation over 0.6. Clearly, these two sets of conditions (right-only variability on one hand and left-and-both sides variable on the other) are being treated differently by subjects.
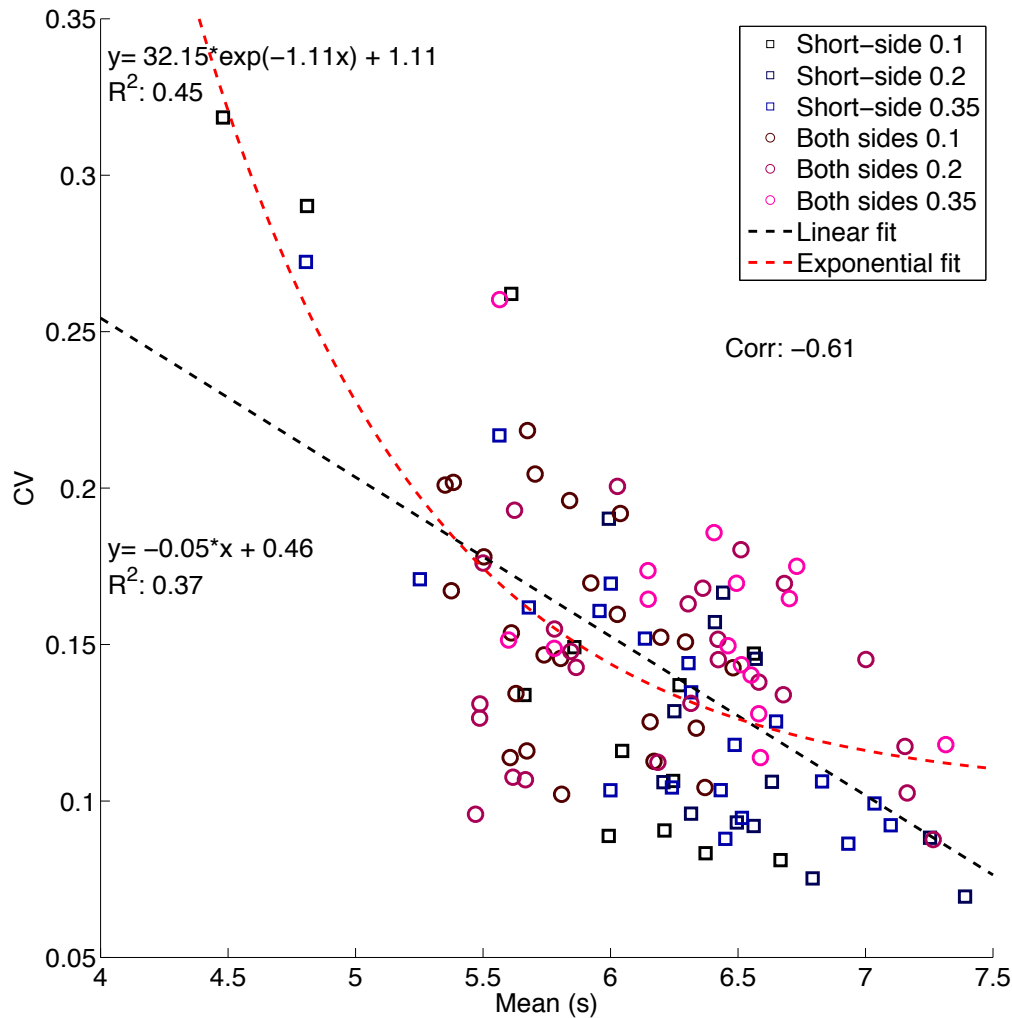
Figure 29. Mean-CV pairings for all conditions with short or both feed latencies variable. These data suggest that CV may not be independent of the interval being timed but rather negatively correlated with it in this range of values. The fit of the exponential function appears to account for more of the data with an R^2 of 0.45 as opposed to 0.37 for the linear.

Discussion

Primary Result

The primary result from this experiment is clear: adding extrinsic variability does not increase the variability in switch latencies, as predicted by popular theories of interval timing; rather it decreases it. Increasing the variability of the durations being timed increases the variability of the subject's record of durations

experienced yet one does not see an increase the variability of timing behavior, as illustrated in Figure 30. In order to account for this result, models of timing must separately represent exogenous and endogenous variability.
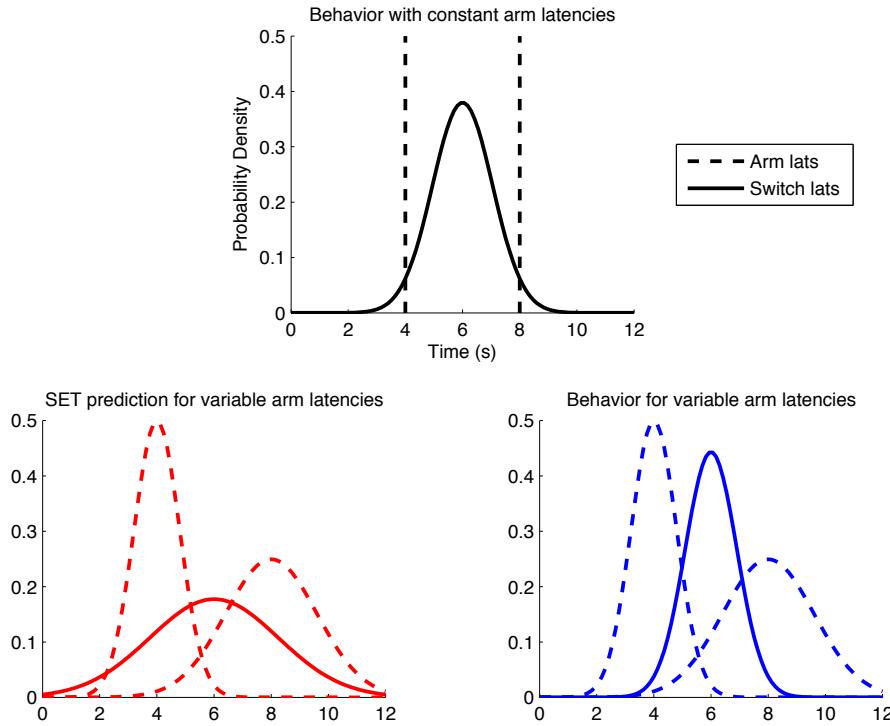
Figure 30. The top panel shows an idealized version of the timing behavior from the constant 4s vs 8s arm latency condition from Experiment 3. The CV shown is the median CV displayed in that data set: CV≈0.175. The bottom two panels show the behavior SET predicts for variable arm latencies and an idealized version of the behavior we observed (in this example, the arm latency CV is 0.2). The CV in timing behavior displayed for SET is the sum of the CV in the top panel (typical observed CV from constant arm latencies) and the arm latency CV (CV≈0.375). Finally, the idealized version of the behavior actually observed is the median CV observed from all "both variable" sessions. Note that the failure of SET's prediction does not depend on any of these particular choices: The CV predicted by SET is considerably higher than any CV observed in this entire experiment

This task is doubly difficult when we take into account the fact that subjects make this distinction even when both arm latencies are given the same CV because the sum of two Gaussians, one with mean zero, is another Gaussian (with variance equal to the sum of the two component variances). Since timing behavior is

Gaussian-distributed, the addition of Gaussian-distributed variability to the events in subjects' environment should not change the form of the distribution of perceived durations. Nevertheless, this experiment suggests that subjects are able to dissociate endogenous from exogenous variability in exactly this case. The two most obvious accounts of how they might be doing this are unsatisfactory for different reasons.

One possible explanation of how subjects learn their timing variability is that subjects are comparing their timing to a different gold standard than trial duration. Because of the sparseness of their environment, the only plausible external gold standard would be the dark/light cycle in their text boxes. This is problematic for two reasons: First, the dark/light cycle does not seem to be a good gold standard in the wild because it fluctuates greatly with the season. Second, the timescale is 7,200 times that of the intervals they are timing in our experiment (light/dark changes every 12 hours). One would imagine that there might be practical difficulties with using a temporal gold standard nearly four orders of magnitude larger than the intervals being timed.

It therefore seems more plausible that subjects estimate their endogenous variability by taking multiple simultaneous measurements from internal timing mechanisms. Many popular theories of timing posit some kind of internal timing mechanism. Setting aside the details of how this timing mechanism works (see General Discussion section), taking readings from concurrently running timing mechanisms would allow subjects to estimate their endogenous timing variability.

This is a straight-forward solution to the problem of estimating endogenous variability and is therefore a strong candidate for further testing.

Secondary Results: Some Puzzles

The fact that timing means drift away from optimality as CVs are increased (at least in "both-variable" conditions, Figure 28) is another indicator that subjects are not behaving as SET and other popular theories would predict: simply optimizing their behavior based on the perceived stimulus durations. If they were, they would not be optimizing well. This is not a satisfying account of their behavior, however, because their behavior does not simply become more variable, it steadily drifts in the opposite direction from optimality without becoming more diffuse. Any viable model of this behavior should account for the precise but inaccurate nature of their responding.

The result that CVs and means were negatively correlated when the short latency was variable (whether or not the long latency was variable) suggests that in some cases we may be seeing decreasing variation rather than constant variation. The wildly surprising aspect of this is that mean and CV are negatively correlated. While this result is deeply interesting for the psychophysics of timing, it unfortunately also does not paint a clear picture of how animals might be separating exogenous from endogenous variability.

Conclusion

The question of how subjects differentiate their timing variability from external variability is perhaps less conceptually important than the fact that they

did make this distinction. This echoes findings in previous experiments presented in this dissertation showing that mice are not simply employing low-level hill-climbing strategies but rather have a rich representation of the problems they are attempting to solve. These subjects made a clear distinction between the variability in their timing system and the variability of the durations of the events they are timing. In a very real way, mice have a concept of objective time as opposed to their perceived time: an allocentric temporal encoding. The mouse's timing mechanism is built under the assumption that when an interval elapses and no mouse is around to observe it, it still exists.

General Discussion

We have established four main results:

1. Subjects shift their timing behavior abruptly in response to a change in the probability of a short trial.

2. The number of trials before this behavioral shift is predicted by the information-theoretic Kullback-Leibler divergence.

3. Timing CVs are not constant and can be lowered by making the temporal discrimination task more difficult (squeezing the "temporal goalposts").

4. Subjects treat additional exogenous timing variability differently than the endogenous variability natural to their timing mechanism.

All four of these results suggest that animal timing and decision behavior be treated in a conceptually better-founded way than has been done thus far. All four suggest that timing and decision behavior is better understood as a system that is solving a specific set of information-theoretic problems: What is the distribution of the events I am engaging with? How is it changing? How can I make the best use of the information available to me? I will now briefly summarize and expand on the basic results before moving on to future directions.

Summary of results

Result 1: behavioral shifts are abrupt

We have seen in Experiment 1 that subjects shift their timing behavior abruptly, not gradually. Since it has been shown that subjects optimize their timing behavior with respect to the probability of a short trial (Balci *et al.*, 2009) we constrain the possible methods by which they perform this optimization by noting that subjects change their behavior all at once, rather than gradually. Specifically, we rule out a wide class of differential reinforcement methods, since many subjects shift before they have missed even a single reinforcer. If subjects were simply noting which strategies for switching yielded more reinforcers, this would be impossible.

This result suggests that subjects are internally representing probability and risk to keep track of what switch strategies are unnecessarily risky, even when they have not yet led them astray. One does not need to drive off the edge of the road several times to develop the behavior of driving comfortably away from the edge of the road.

Result 2: KL-Divergence predicts behavioral shift latency

Experiment 2 showed that Kullback-Leibler divergence predicts behavioral shift latency (the number of trials from the change in the probability of a short trial to a change in timing behavior). More specifically, transitions from 0.5 probability of a short trial to 0.1 probability causes a behavioral shift fewer trials than a transition in the opposite direction *using the very same probabilities.* Associative theories of learning do not predict this result. I have already argued this is the case based on *a priori* mathematical considerations. To further demonstrate this I will now argue the same point via numerical

simulation. Assume a model has a memory of the last N trials as either being long (1) or short (0) and its output is driven by a geometrically-discounted summary of this memory as in

$$M = \frac{\sum_{i=0}^{N} \delta^i * t_{N-1}}{\sum_{i=0}^{N} \delta^i}$$

where t_n is the record of the nth trial. Note that when we change the probability of a long trail (the expected value of t_i) the expected value of the summary M changes as more trials at the new probability value are experienced. The changes are dramatic at first because the more recent trials are more heavily weighted. Changes in M do not have the same asymmetric properties as the Kullback-Leibler divergence. If we consider a memory full of N trials generated with probability p_1 of a long trial into which we begin introducing trials generated with probability p_2, M will shift from p_1 to p_2. The changes in M will be quick at first and then gradual because the most recent trials are the most heavily weighted. If we compare this to a change from p_2 to p_1, we can see that the magnitude of the change expected after seeing k trials at the new probability value is the same going from p_1 to p_2 or vice versa, as in Figure 31. This means that he convergence of M to the actual value is symmetrical and therefore behavior based on M, as predicted by typical associative theories of learning, should also be expected to be symmetrical. On the other hand, Kullback-Keibler divergence, which is the speed of convergence of a Bayesian model (see below), is asymmetrical, in line with our data.
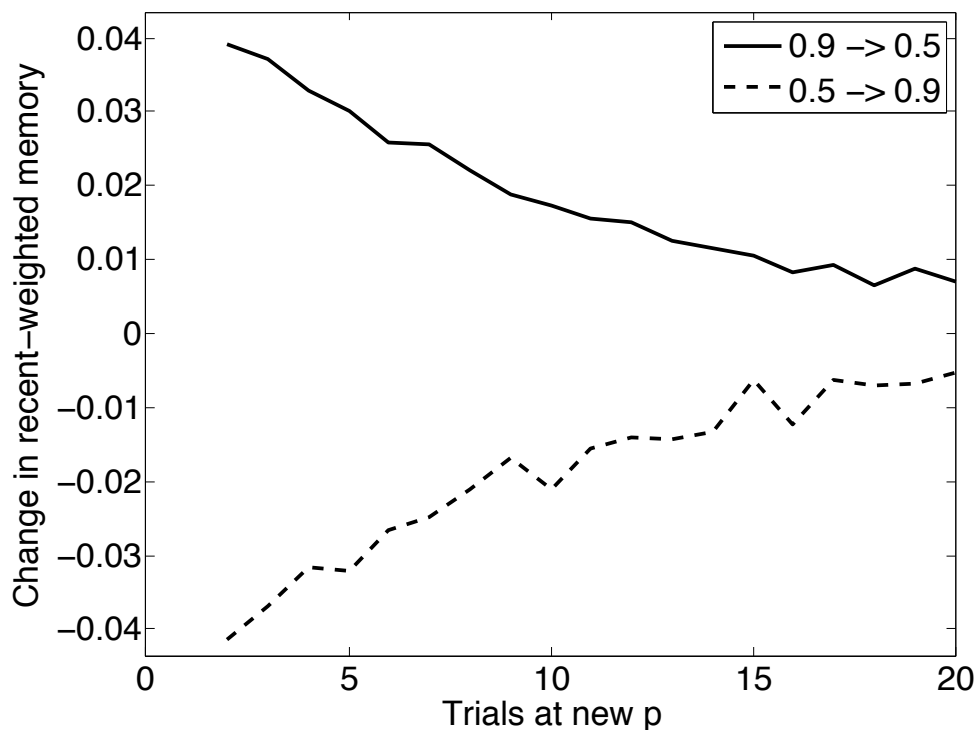
Figure 31. Expected change in M after it has experienced a given number of trials at the new probability. The change (apart from sign) is symmetrical whether going from probability 0.9 to 0.5 or vice versa. This is different from the behavior of the Kullback-Leibler divergence and from our data.

Beyond than arguing against associative theories of learning this result suggests that an information-theoretic framework is necessary to understand the change detection that our subjects are performing. This asymmetry in change detection is highly counterintuitive but is an elementary prediction of a basic examination of this problem in information-theoretic terms. That is to say, if we understand subjects to be evaluating the probability that the environment has changed from one distribution to another via anything like a Bayes factor or likelihood ratio, these asymmetries naturally emerge.

Take as a toy example a model in which a subject is deciding whether an incoming stream of data are being generated by distribution A or distribution B when they are in fact being generated by distribution A. For simplicity's sake, let's assume that these are the only two possibilities and the subject knows it. All plausible methods of

comparing these two hypotheses are at their core a ratio of the likelihoods of the data given distribution A and the data given distribution B. For our toy example, let us take a toy decision criterion: declare that A is the true distribution iff.

$$k < \frac{P(D|A)}{P(D|B)}$$

for some constant criterion k. Taking logs of both sides, this is equivalent to

$$\log_2 k < \log_2\left(\frac{P(D|A)}{P(D|B)}\right)$$

The right-hand side of this, the quantity we are comparing to a constant in order to make our decision changes as each new datum comes in. From a god's-eye view, the expectation for how much this quantity will change with the next datum is the log of the ratio of the two distributions (evaluated over their entire support) weighted by the probability of getting each value:

$$\Delta ratio = \sum_i \log_2\left(\frac{P(i|A)}{P(i|B)}\right) \cdot P(i|A)$$

$$= \sum_i \log_2\left(\frac{A(i)}{B(i)}\right) \cdot A(i)$$

$$= D_{KL}(A \parallel B)$$

where i ranges over the entire support.

So, we see that in this simple example, the KL divergence is the amount we expect the log-likelihood ratio to change each trial. If, in turn, subjects' decisions are driven by the change in log-likelihood ratio, we naturally expect the asymmetries we have found in the speed with which they shift their behavior. Such calculations, however, are only possible in models that explicitly represent probabilities and are capable of performing symbolic computations. If the subject is not detecting changes

but rather simply operating on a best estimate of the current probability based on the frequency of long trials in the last N, we would not expect to see such asymmetries.

An information-theoretic framework also naturally nests to incorporate higher-level effects not considered here. For example, one might use the distribution of trials-experienced-before-change to predict how far off the next change is likely to be as in Gallistel, Mark, King, & Latham (2001) or use the distribution of distributions experienced to more as seen in quickly categorize a new condition. In procedures more sensitive to the location of a behavioral shift, accounts of this kind also provide an optimal standard to which subject performance can be compared.

Result 3: Timing CVs can be lowered when the timing discrimination is difficult

When we lowered the long feed latency, thereby squeezing the "temporal goalposts" to make the task more difficult, subjects increased the accuracy of their timing relative to the interval being timed. This is a violation of Weber's law, which has been taken to be a fundamental property of interval timing (Gibbon & Church, 1984). It is possible that Weber's law still holds for timing behavior in sufficiently well-motivated mice but this has yet to be tested. This could easily be done using a paradigm similar to Experiment 3 but varying the short and long feed latencies over a wider range of values. Experiment 3 has given us data on the CV for one difficult discrimination (4s vs. 6s): a test for Weber's law would be to see whether we see the same CV for equally difficult discriminations with different absolute magnitudes (8s vs. 12s, 16s vs. 24s, etc.).

Result 4: Subjects distinguish exogenous and endogenous sources of variability

Subjects responded differently to exogenous variability than they did to the variability endogenous to their timing mechanism. As discussed above, this cannot be accounted for by most currently popular theories of timing. This is because these popular theories typically do not represent uncertainty explicitly, which is a natural feature of information processing theories. The results of this experiment show that mice are sensitive to such uncertainty; in many conditions decreasing their timing variability when the variability in their stimuli increases. This result is interesting in itself but perhaps most interesting because it demands models of animal behavior that are more sophisticated than has historically been popular.

As argued in the Discussion section of this experiment, these data suggest that subjects estimate their timing variability not in reference to some temporal gold standard. This could be achieved by simultaneously taking readings from several independent timing mechanisms. An initially appealing option is to model this as simply being several copies of the accumulator model that underpins SET and many other models of animal timing. This account, however, is ultimately unattractive for theoretical reasons: it compounds theoretical problems already facing models such a SET. Accumulator models such as the one in question essentially function as stopwatches: when an interval starts, the subject begins accumulating pulses and when it ends, the subject takes a reading. This requires that subjects know when an interval begins. This may not always be possible when the subject does not know in advance which cues predict the event in question. When a new event happens, it seems difficult to give an account of how many

accumulators a subject would need to start and how it would know when to reset them.

If one needs to keep track of several overlapping durations, as no doubt animals must do in the wild, the problem of how many timers to set grows exponentially with the number of possibly interconnected events the subject is keeping track of. Worse yet, subjects would be incapable of using a past event as a temporal marker if they didn't realize it was a temporal marker at the time: if you learn at lunch that breakfast time predicts dinner time then you haven't set the appropriate timer and therefore are completely incapable of predicting when dinner will be served. Applying this kind of knowledge retroactively can be seen in so-called "backward conditioning" (Arcediano & Miller, 2002).

A slight modification of this is to use constantly running timers rather than accumulators. This is like the difference between using the system clock on a computer versus using a stopwatch. A system clock can be used as a stopwatch by simply recording the time at the beginning of an interval, recording the time at the end of an interval and then taking the difference of the two to find the duration of the interval. Moreover, by simply keeping a record of the times of various events, one can recover arbitrary temporal differences at any point in the future. This means that subjects do not need to start and stop a wide variety of timers in order to keep track of multiple overlapping intervals. A single timestamp record, combined with the ability to perform calculations with timestamps, allows subjects to recover temporal differences as needed.

If subjects had a modest number of such timers, they could estimate their timing variability without reference to a temporal gold standard. Our data do not speak to whether each timer would need to have its own timestamp record or whether a single record could be kept that would be some kind of average of the readings from the clocks. The former would allow subjects to calculate how their variability changed over time whereas the latter would only allow for an estimate of the current timing variability.

Doubtless, one objection to this account is that it does not follow Weber's law. This is because it has increasing relative temporal precision as time increases. If one is reading start and stop times off an internal clock, timing a 9s interval one day into one's life means taking the difference of a start measurement of 86,400s and a stop measurement of 86,409s. A month into one's life, timing the same interval means taking a start measurement of 2,592,000s and a stop measurement of 2,592,009s. There are very few systems in nature in which absolute precision does not decrease with the quantity being measured; in which the difference between 1 and 2 units has as big an effect as the difference between 101 and 102 units.

The system clocks theory does appear to necessitate both constant timing precision and a memory for storing timestamps beyond what is typically assumed to be available to subjects. In reply to these facts I can only repeat, as I pointed out in the introduction, that the physical bases of memory and timing are not well understood and therefore should not limit the theory of the computations underlying behavior. In this sense, it is in the spirit of information processing theories of behavior: storing and computing over raw information as necessary. It is

uncontroversial that training from many trials in the past can affect behavior, as in rapid reacquisition of a learned behavior, in which an association that has been learned and then trained-out through extinction is then relearned more rapidly than in the initial learning phase. This relearning is clearly sped up by the initial learning phase despite the fact that this happened many trials ago. Under the system clocks theory, this is because past examples are remembered whereas in popular behavioral theories such as memory trace models, this is because of the residual effect of previous trials on one or two variables.

For this reason, it is possible that the system clocks theory combined with information processing calculations could better handle patterns of behavior such as spontaneous recovery that are awkward to fit into popular theories that change parameters incrementally, as discussed above. A still-unpublished but elegant and insightful development of this line of thinking has been produced (Wilkes, 2013) that predicts many of the results (including counter-intuitive ones) of classical conditioning from basic information-theoretic principles.

Take-home messages

I will now attempt to take these results as a group and place them back in the context I originally put them in. I will move from the context of theories of animal behavior to the broader context of decision under uncertainty.

We are observing an information-processing machine

I argue that these results show that animals even in as simple as mice, behavior is driven by sophisticated information-theoretic calculations. They are

fundamentally not engaged in guess-and-check, hill-climbing or other simple strategies. These calculations represent abstract properties of the environment, such as the probability of an event, and calculate near-optimal responses using those representations.

I would like to stress again that I don't believe that animals are engaged in the same type of mental activity a freshman probability student is engaged in when she performs a calculation for an exam. Speaking of mice calculating probabilities can often lead to this kind of picture that I believe is rightfully regarded as incorrect. Computers, however, also aren't engaged in the same kind of mental activity as that freshman and yet we are happy to say that they are capable of calculating a probability or detecting a change.

In a similar way, many researchers are comfortable talking about the visual system performing complex calculations (performing a Fourier decomposition on shapes to transform a retinal display into blips that the brain will interpret as shapes). I believe that it is in this sense that the brain is also performing complex Bayesian analyses on the probabilities and events we encounter. Gut feelings don't just come from nowhere: they are the product of deep and meaningful analysis.

At the same time, I do not wish to suggest that the non-verbal reasoning is somehow perfect. Doubtless, animals make mistakes in their decisions under uncertainty. For example, subjects drifted away from optimality as exogenous variability increased in Experiment 4. I do, however, think the results above underscore that their reasoning is an information-processing machine that works under uncertainty and with noisy signals and is therefore best understood in an

information-processing framework. This view is not so outlandish when one considers the fact Bayesian inference is merely the formalization of the complex problems animals have faced in the wild since time immemorial: making decisions based on limited information with varying degrees of reliability.

I have therefore focused on questions about how a system could calculate appropriate behavioral responses and what inputs and variables it would need to do so. I have shown that the above data cannot be explained by theories that have primarily aimed to explain how behavior could be timed to coincide with temporal regularities in a stimulus. Rather, the data support theories capable of much more calculation.

Animals calculating is nothing new

These results are certainly not the first to suggest that animals are capable of calculating responses rather than locally tweaking their behavior. Two examples are particularly salient in the given context. One example is the time-left experiment of Gibbon and Church (Gibbon & Church, 1981). In this experiment, rats were given a choice between a standard fixed-interval of 15s to reinforcement at one lever and the time left to reinforcement at an elapsing comparison interval of 60s. When the fixed interval was initiated early in the course of the comparison interval, subjects greatly preferred responding at that lever. When the fixed interval was initiated late in the course of the comparison interval, subject greatly preferred the comparison interval. When the fixed interval was initiated halfway through the comparison interval, subjects were roughly indifferent. Gibbon and Church interpreted this result as preference being driven by which lever would be ready to deliver a

reinforcer first. This decision appears to require that the subjects be able to subtract times so as to calculate which lever will be armed sooner.

The second example is the fact that the number of trials required to acquire a conditioned response in Pavlovian autoshaping (for a variety of species and preparations) depends linearly on the ratio of the time the subject is exposed to the experimental apparatus and the CS-US interval. To explain this and similar informationally-driven results, Gallistel and Gibbon (2000) posited Rate Estimation Theory (RET). RET relies on the ability of the subject to perform calculations over times of occurrence and is motivated by considering at what point the unconditioned stimulus becomes a significantly better predictor of the unconditioned stimulus than merely being in the test environment. It explicitly computes the predictivity of the conditioned stimulus, which is the intellectual foundation for the information-processing framework advocated here.

Not picking sides: Gigerenzer and Kahneman

I do not think the results of this dissertation contribute to a view of decision in line with Kahneman's but neither do they support a view in line with Gigerenzer's. I think they subvert the classical idea that non-verbal reasoning (the kind of which even animals are capable) is simpler than the conscious, verbally-based reasoning that we engage in when we make a pros and cons list in buying a new stereo or taking a new job. In a broad sense, this aligns with the view that our instincts are smarter than previously thought and to be trusted. At the same time, it goes against Gigerenzer's view of the non-conscious mind as useful because of its great simplicity, which I see as a kind of "noble savage" view of the non-conscious

mind. I see no evidence here for an adaptive toolbox of highly specialized heuristics. Rather, I see evidence for a sophisticated program that is nevertheless robust because it deals with unreliable and incomplete information and explicitly takes that into account.

The reason for trusting gut feelings is not merely their simplicity but the fact that they are the product of sophisticated yet still robust calculations. The work in this dissertation does not speak to the biases and cognitive traps that appear to plague conscious decision-making except to underscore what a puzzle they are given that non-conscious reasoning can be so accurate and robust. The way in which our intuitions interact with our conscious mind seems to me a fertile field for exploration. These issues touch on deep questions about what consciousness is and what it is for and therefore what it is to be human.

I believe the results above are most important because of the deep questions they suggest looking forward: If even mice are capable of near-optimal Bayesian inference, how much more is the human mind capable of? Why are we so often led astray by the traps and biases catalogued by Kahneman and Tversky and colleagues? What makes our reasoning different from that of animals? If animals are capable of this type of calculation, how different are we really?

References

Arcediano, F., & Miller, R. R. (2002). Some constraints for models of timing: A temporal coding hypothesis perspective. *Learning and motivation, 33*(1), 105-123.

Armus, H. L., Montgomery, A. R., & Jellison, J. L. (2006). Discrimination Learning in Paramecia. *The Psychological Record, 56*, 489-498.

Balci, F. (2007). *Risk assessment in mice and men.* Rutgers University, New Brunswick, NJ.

Balci, F., Freestone, D., & Gallistel, C. R. (2009). Risk assessment in man and mouse. *Proceedings of the National Academy of Sciences, 106*(7), 2459–2463.

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*(9), 1214-1221.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in Visual Working Memory: Using Statistical Regularities to Form More Efficient Memory Representations. *Journal of Experimental Psychology: General, 138*(4), 487-502.

Briscoe, E., & Feldman, J. (2011). Conceptual complexity and the bias/variance tradeoff. *Cognition, 118*(2-16).

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience, 6*(10), 755-765.

Chater, N., Tenenbaum, J. B., & Yuille, A. (2006). Probabilistic Models of Cognition: Conceptual Foundations. *Trends in Cognitive Sciences, 10*(7), 287-291.

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*(4), 698-707.

Courville, A. C., Daw, N. D., & Touretzky, D. S. (2006). Bayesian theories of conditioning in a changing world. *Trends in Cognitive Sciences, 10*(7), 294-300.

Dayan, P. (2012). Robust Neural Decision-Making. In *Evolving the Mechanisms of Decision Making: Toward a Darwinian Decision Theory.* Frankfurt: Ernst Strüngmann Forum.

Gallistel, C., King, A., Daniel, A., Freestone, D., Papachristos, E., Balci, F., Kheifets, A., Zhang, J., Su, X., & Schiff, G. (2010). Screening for learning and memory mutations: A new approach. *Xin li xue bao. Acta psychologica Sinica, 42*(1), 138.

Gallistel, C., King, A., & McDonald, R. (2004). Sources of variability and systematic error in mouse timing behavior. *Journal of Experimental Psychology: Animal Behavior Processes, 30*(1), 3.

Gallistel, C., Krishan, M., Liu, Y., Miller, R., & Latham, P. (2013). The Perception of Probability. *Psychological Review, (in press)*.

Gallistel, C., Mark, T. A., King, A. P., & Latham, P. (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*(4), 354.

Gallistel, C. R., Balci, F., Freestone, D., Kheifets, A., & King, A. P. (2013). Automated, Quantitative Cognitive/Behavioral Screening of Mice: For Genetics, Pharmacology, Animal Cognition and Undergraduate Instruction. *The journal of Visualized Experiments, in press*.

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review, 107*(2), 289.

Gallistel, C. R., & King, A. P. (2009). *Memory and the computational brain: why cognitive science will transform neuroscience* (Vol. 3): Wiley-Blackwell.

Getty, D. J. (1975). Discrimination of short temporal intervals: A comparison of two models. *Perception & Psychophysics, 18*(1), 1-8.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84*(3), 279.

Gibbon, J. (1999). Multiple time scales is well named. *Journal of The Experimental Analysis of Behavior, 71*(2), 272.

Gibbon, J., & Church, R. M. (1981). Time left: linear versus logarithmic subjective time. *Journal of Experimental Psychology: Animal Behavior Processes, 7*(2), 87.

Gibbon, J., & Church, R. M. (1984). *26 Sources of variance in an information processing theory of timing.* Paper presented at the Animal Cognition: Proceedings of the Harry Frank Guggenheim Conference, June 2-4, 1982 [held at Columbia University].

Gibbon, J., Malapani, C., Dale, C., & Gallistel, C. (1997). Toward a neurobiology of temporal cognition: advances and challenges. *Current Opinion in Neurobiology, 7*(2), 170-184.

Gigerenzer, G. (2004). Fast and frugal heuristics: The tools of bounded rationality. In D. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 62-88).

Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*: Penguin.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology, 62*, 451-482.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological Review, 103*(4), 650.

Gigerenzer, G., & Selten, R. (2002). *Bounded rationality: The adaptative toolbox*: Mit Press.

Gladwell, M. (2007). *Blink: The power of thinking without thinking*. New York, NY: Back bay Books/Little, Brown and Company.

Glöckner, A., Betsch, T., & Schindler, N. (2010). Coherence shifts in probabilistic inference tasks. *Journal of Behavioral Decision Making, 23*(5), 439-462.

Goldstein, D. G., & Gigerenzer, G. (2009). Fast and frugal forecasting. *International Journal of Forecasting, 25*(4), 760-772.

Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: exploring representations and inductive biases. *Trends in Cognitive Sciences, 14*, 357-364.

Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal Predictions in Everyday Cognition. *Psychological Science, 17*(9), 767-773.

Griffiths, T. L., & Tenenbaum, J. B. (2011). Predicting the future as Bayesian inference: People combine prior knowledge with observations when estimating duration and extent. *Journal of Experimental Psychology: General*.

Grondin, S., & Killeen, P. R. (2009). Tracking time with song and count: Different Weber functions for musicians and nonmusicians. *Attention, Perception, & Psychophysics, 71*(7), 1649-1654.

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences, 13*(12), 517-523.

Hilbig, B. E. (2010). Reconsidering "evidence" for fast-and-frugal heuristics. *Psychonomic Bulletin & Review, 17*(6), 923-930.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263-291.

Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007). Learning Overhypotheses with Hierarchical Bayesian Models. *Developmental Science, 10*(3), 307-321.

Kemp, C., & Tenenbaum, J. B. (2008). The Discovery of Structureal Form. *Proceedings of the National Academy of Sciences, 105*(31), 10687-10692.

Kheifets, A., & Gallistel, C. R. (2012). Mice Take Calculated Risks. *Proceedings of the National Academy of Sciences*.

Killeen, P. R., & Fetterman, J. G. (1988). A behavioral theory of timing. *Psychological Review, 95*(2), 274.

Lewis, P., & Miall, R. (2009). The precision of temporal judgement: milliseconds, many minutes, and beyond. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1525), 1897-1905.

Machado, A. (1997). Learning the temporal dynamics of behavior. *Psychological Review, 104*(2), 241.

Maloney, L. T., & Mamassian, P. (2009). Bayesian decision theory as a model of human visual perception: Testing Bayesian transfer. *Visual Neuroscience*(26), 147.

McLeod, P., & Dienes, Z. (1996). Do fielders know where to go to catch the ball or only how to get there? *Journal of Experimental Psychology: Human Perception and Performance, 22*(3), 531.

Meck, W. H. (2003). *Functional and neural mechanisms of interval timing*: CRC Press.

Montemayor, C. (2012). *Minding Time: A Philosophical and Theoretical Approach to the Psychology of Time* (Vol. 5): Brill.

Newell, B. R. (2005). Re-visions of rationality? *Trends in Cognitive Sciences, 9*(1), 11-15.

Oppenheimer, D. M. (2003). Not so fast! (and not so frugal!): Rethinking the recognition heuristic. *Cognition, 90*(1), B1-B9.

Rieke, F., Warland, D., De Ruyter van Steveninck, R., & Bialek, W. (1997). Exploring the neural code: The MIT press.

Staddon, J., & Higa, J. (1999). TIME AND MEMORY: TOWARDS A PACEMAKER-FREE THEORY OF INTERVAL TIMING. *Journal of The Experimental Analysis of Behavior, 71*(2), 215-251.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine learning, 3*(1), 9-44.

Tenenbaum, J. B. (1999a). *A Bayesian Framework for Concept Learning.* Unpublished Thesis, Massachusetts Institute of Technology, Cambridge, MA.

Tenenbaum, J. B. (1999b). Bayesian Modeling of Human Concept Learning. In M. Kearns, S. Solla & D. Cohn (Eds.), *Advances in Neural Information Processing Systems* (Vol. 11, pp. 59-65). Cambridge, MA: MIT Press.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. (2011). How to Grow a Mind: Statistics, Structure, and Abstraction. *Science, 331*(March), 1279-1285.

Trommershäuser, J., Gepshtein, S., Maloney, L. T., Landy, M. S., & Banks, M. S. (2005). Optimal compensation for changes in task-relevant movement variability. *The Journal of Neuroscience, 25*(31), 7169-7178.

Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2008). Decision making, movement planning and statistical decision theory. *Trends in Cognitive Sciences, 12*(8), 291-297.

Trommershäuser, J., Maloney, L. T., & Sandy, M. S. (2003). Statistical decision theory and tradeoffs in motor response. *Spatial Vision, 16*, 255-275.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*(4), 293.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of business*, S251-S278.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty, 5*(4), 297-323.

Van Hamme, L. J., & Wasserman, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning and motivation, 25*(2), 127-151.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences, 7*(11), 483-488.

Wilkes, J. (2013). Inference, Invariance, and Representation: Weber's Law and the Foundations of Cognitive Science. *Unpublished*.