

# Learning to Detect Change

Ye Li<sup>1,\*</sup>, Cade Massey<sup>2</sup>, and George Wu<sup>3</sup>

*\* Corresponding author*

<sup>1</sup>*Columbia University, Center for Decision Sciences, 3022 Broadway Ave., New York, NY 10027,  
yl2629@columbia.edu, 847.666.5353*

<sup>2</sup>*Yale University, School of Management, 135 Prospect St., New Haven, CT 06511, cade.massey@yale.edu*

<sup>3</sup>*University of Chicago, Booth School of Business, 5807 S. Woodlawn Ave., Chicago, IL 60637,  
wu@chicagobooth.edu*

Affiliation

## Abstract

People, across a wide range of personal and professional domains, need to accurately detect change. Previous research has documented a systematic pattern of over- and underreaction to signals of change due to system neglect, or the tendency to overweigh signals and underweigh the system producing the signals. We investigate whether people can improve at change detection with experience. We find that the system-neglect pattern persists, but that learning varies across environments—participants showed reliable improvement in some conditions and virtually none in others. We explain this differential learning by formally characterizing environments in terms of how much they (1) provide consistent feedback, and (2) tolerate suboptimal behavior. Counter-intuitively, tolerating errors may help learning because strong incentives draw attention to performance and away from a deeper understanding of complex tasks. We discuss extensions to other stochastic environments and implications for organizations, and introduce a simple heuristic to assist in change detection.

**Key Words:** Change-point detection, regime shift, learning, overreaction, underreaction, Bayesian updating, simple heuristics, probability judgment.

## INTRODUCTION

The need to detect change accurately is a common problem for people in a wide range of domains, from business and politics to social relations and sports. The canonical academic example involves monitoring quality levels in a manufacturing process (Deming, 1975; Rubin & Girshick, 1952; Shewhart, 1939), but recent work has broadened this paradigm beyond operations research. Researchers in finance have used it to explain well-documented patterns of over- and underreaction to news in stock prices (Barberis, Shleifer, & Vishny, 1998; Brav & Heaton, 2002), and economists have used change-point models to describe the challenges central bankers face in setting interest rates (Ball, 1995; Blinder & Morgan, 2005). Even further afield, change detection is important to retailers assessing changes in consumer taste (Fader & Lattin, 1993), corporate strategists monitoring technological trends (Grove, 1996), politicians tracking voter sentiment (Bowler & Donovan, 1994), and individuals keeping track of their health (Steineck et al., 2002) or their romantic partner's commitment (Sprecher, 1999). Indeed, the need to detect change accurately is ubiquitous, and it is therefore critical to understand the behavioral patterns involved in doing so.

These examples highlight the challenge of successfully identifying change: One must infer the true state or “regime” from unreliable signals while balancing the costs of underreacting (failing to realize change has occurred) against the costs of overreacting (believing change has occurred when in fact it has not). For example, an investor needs to recognize when financial markets change from a “bear” to a “bull” market. Economic indicators are, at best, imprecise signals, with informativeness varying across indicators and over time. Underreacting to signals of change would mean foregoing the chance to buy stocks at their lowest prices, whereas overreacting would mean acquiring shares that are still declining.

A number of psychologists have investigated how successfully individuals navigate this task (e.g., Barry & Pitz, 1979; Brown & Steyvers, 2009; Chinnis & Peterson, 1968, 1970; Estes, 1984; Rapoport, Stein, & Burkheimer, 1979; Robinson, 1964; Theios, Brelsford, & Ryan, 1971). A theme in this literature is that individuals respond to system parameters, but only partially. As Chinnis and Peterson (1968) stated: “subjects, while sensitive to the difference in diagnostic value of the data in the two conditions, were not adequately sensitive” (p. 625). Massey and Wu (2005a) recently shed light on this problem, proposing the *system-neglect hypothesis*: People react primarily to the signals they observe and secondarily to the system that produced the signal. In their experiments, participants were exposed to signals generated by a number of different systems in which diagnosticity (i.e., the precision of the signal) and transition probability (i.e., the stability of the system) were varied. In our investor example, diagnosticity corresponds to the informativeness of the market indicators and transition probability corresponds to the historical rate of market vacillation. Their experiments revealed a behavioral pattern consistent with system neglect: Underreaction was most common in unstable systems with precise signals, whereas overreaction was most prevalent in stable systems with noisy signals. Kremer, Moritz, and Siemsen (2011) found additional support for system neglect in a time-series environment with continuous change and continuous signals.

A common critique of behavioral decision research is that participants engage in relatively novel tasks in unfamiliar environments, without sufficient opportunity to learn (Coursey, Hovis, & Schulze, 1987; List, 2003). In Massey and Wu (2005a) (hereafter: MW), for example, participants changed system diagnosticity and transition probability after each of 18 trials. On the one hand, this arrangement enhanced the salience of each system's variables, increasing the likelihood that participants would give them sufficient attention. On the other hand, their design may

have hindered participants’ ability to appropriately adjust to these dimensions by minimizing their opportunity to learn about a particular system and raises questions about the robustness of the system-neglect hypothesis. Does the pattern of over- and underreaction observed in MW persist in the face of learning? Or is the phenomenon only observed in those inexperienced with the task and therefore less applicable to real-world settings where people have sufficient opportunities to learn? The present paper aims to fill this gap in knowledge.

The rest of this paper is organized as follows. We begin by describing the statistical process used in our experiment and outline our hypotheses. We then present the experimental design and results, in particular examining the relationship between decision environments and learning. Next we propose a simple, psychologically plausible decision heuristic as a way to model participant behavior as well as to ease the learning process. We conclude by discussing implications for learning in other stochastic environments.

## THEORY

In this section, we introduce the design of our experiment, review the system-neglect hypothesis predictions for detecting changes in this statistical process, and propose new hypotheses on how learning to detect change depends on certain characteristics of the learning environment.

### *Terminology*

Before we begin, we introduce key terms used throughout the paper. First, the *system* is the random process that generates binary *signals* (red or blue balls) in each of the 10 *periods* that make up a *trial*. A system is characterized by two system variables: *diagnosticity*, or the informativeness of the signals it generates, and *transition probability*, or the instability of the system. The systems are dynamic in the sense that they can generate signals using two sets of probabilities, each of which we call a *regime*. Finally, a *learning environment* consists of the system along with any task parameters such as payment and feedback characteristics.

### *Design*

Our experimental paradigm largely mirrors the one used in MW. In each trial of 10 periods, the system begins in the red regime but has a transition probability  $q$  of switching to the blue regime in any period  $i$  (including in the first period before any signal is drawn). If the system switches to the red regime, it does not switch back. That is, the blue regime is an absorbing state.

The system generates either a red or blue signal in each period. A red signal is generated by the red regime with probability  $p_R > .5$  and by the blue regime with probability  $p_B < .5$ . The probabilities were symmetric in our experiment (i.e.,  $p_R = 1 - p_B$ ), so  $p_R/p_B$  is a measure of the diagnosticity ( $d$ ) of the signal, with larger diagnosticities corresponding to more precise or informative signals. Participants were given all of these system variables and told that their task was to guess which regime generated that period’s signal. In particular, they estimated the probability that the system had switched to the blue regime. Importantly, at the end of each trial, participants received feedback about the true regime in each period of that trial.

We constructed six experimental conditions by crossing three diagnosticity levels ( $d = 1.5, 3,$  and  $9$ ) with two transition probability levels ( $q = .05$  and  $.20$ ). We will refer to these systems qualitatively as noisy ( $d = 1.5$ ) versus somewhat precise ( $3$ ) versus highly precise ( $9$ ) and stable ( $q = .05$ ) versus unstable ( $.20$ ). Our most important deviation from MW is that within each condition, the system variables remained constant across all 20 trials. Although we pre-generated

the 20 random trials of 10 periods each for each condition, participants received the 20 trials in randomized order. The actual series for each trial can be found in Appendix A.

We compensated participants according to a quadratic scoring rule that paid a maximum of \$0.08 (e.g., if a participant indicated a 100% probability that the system was in the blue regime and it in fact was) and a minimum of -\$0.08 (e.g., if a participant indicated a 100% probability that the system was in the blue regime but it was in fact still in the red regime). A quadratic scoring rule theoretically elicits true beliefs for a risk-neutral participant (Brier, 1950). Although it was possible to lose money overall, doing so was extremely unlikely. To control for differences in difficulty across different conditions, we equalized average pay by adding fixed payments to three of the six conditions so that an optimal decision-maker would earn approximately the same amount in each condition (see Table 1).

Optimal responses to the task required application of Bayes' Rule. Let  $B_i = 1$  indicate that the system is in the blue regime in period  $i$ , and let  $b_i = 1$  indicate that a blue signal is observed in that period. If  $H_i = (b_1, \dots, b_i)$  is the history of signals through period  $i$ , the Bayesian posterior odds of a change to the blue regime after observing a history of signals  $H_i$  is:

$$\frac{p_i^b}{1 - p_i^b} = \left( \frac{1 - (1 - q)^i}{(1 - q)^i} \right) \sum_{j=1}^i \frac{q(1 - q)^{j-1}}{1 - (1 - q)^j} d^{\left[ i+1-j - \left( 2 \sum_{k=j}^i b_k \right) \right]} \quad (1)$$

where  $p_i^b$  denotes the probability that the system has switched to the blue regime by period  $i$ . The derivation for Equation (1) is found in Massey and Wu (2005b).

Our experimental setting allowed us to compare individual judgments against the normative standard of Bayesian updating, as we provided participants with all the information necessary to calculate Bayesian responses and hence provide optimal judgments. Therefore, our experiment was designed to test the system-neglect hypothesis by investigating whether individuals update probability judgments in the direction required by Bayesian updating and whether their performance improves with learning.

### THE SYSTEM-NEGLECT HYPOTHESIS

The system-neglect hypothesis posits that participants will be more sensitive to signals than to system variables. This hypothesis extends work by Griffin and Tversky (1992) in stationary environments suggesting that individuals are disproportionately influenced by the strength of evidence (e.g., the effusiveness of a letter of recommendation) at the expense of its weight (e.g., the credibility of the letter writer). This relative sensitivity to strength over weight determines a person's confidence, leading to a pattern of overconfidence when strength is high and weight is low, and underconfidence when strength is low but weight is high. In the context of our dynamic statistical process, the signal (i.e., the sequence of red and blue signals) is the strength, whereas the system variables (i.e., the transition probability,  $q$ , and the diagnosticity,  $d$ ) are the weight. The critical implication of system-neglect is that individuals are more likely to overreact to blue signals in stable systems with noisy signals, and are more likely to underreact in unstable systems with precise signals. However, note that system neglect makes a *relative* prediction and is silent about overall levels of reaction; as such, it is consistent with patterns of only underreaction or only overreaction.

To give a concrete example, consider four systems crossing two levels of diagnosticity,  $d = 1.5$  and  $d = 9$ , with two transition probabilities,  $q = .05$  and  $q = .20$ . Suppose that signals in the first

two periods are both blue. The Bayesian posterior probabilities of a change to the blue regime are  $P(B_2|H_2) = .17$  when  $d = 1.5$  and  $q = .05$  (i.e., the noisy/stable system), but  $P(B_2|H_2) = .92$  when  $d = 9$  and  $q = .20$  (i.e., the precise/unstable system). If individuals give the same response across all four conditions (for example, with a posterior probability of .60), they will overreact when  $d = 1.5$  and  $q = .05$  and underreact when  $d = 9$  and  $q = .20$ . Of course, we do not expect participants to ignore the system variables entirely. However, the system-neglect hypothesis requires that people attend too little to diagnosticity and transition probability and too much to the signals.

### LEARNING HYPOTHESES

Intuition suggests that the hypothesized system-neglect pattern should attenuate over time, with judgments becoming more Bayesian with experience. Although there is little doubt experience *can* lead to learning, it is also clear that it does not always (Brehmer, 1980). Since the system-neglect hypothesis alone provides little guidance on which combinations of diagnosticity and transition probability mitigate or facilitate learning, we turn to the learning literature, using Hogarth (2001) as a starting point. Hogarth identified two variables that affect learning: the quality of feedback and the consequence of errors. That is, learning hinges on whether feedback is informative, timely, and unambiguous, as well as whether inaccurate judgments are penalized appropriately. We take these two dimensions as a starting point for characterizing the conduciveness of environments toward learning.

Mapping feedback quality and the consequence of error to the problem of change detection is not trivial. Whereas the relation between feedback quality and signal diagnosticity is intuitive, the relation between feedback quality and transition probability is not immediately clear. Even less clear is the relation between the consequence of error and either signal diagnosticity or transition probability. Consequently, we look beyond the system variables and consider the learning environment as a whole in order to operationalize feedback quality and the consequence of error.

First, we consider a learning environment's *consistency*, intended to capture Hogarth's quality of feedback. In this research, consistency measures the extent to which optimal behavior on any given experimental trial (one set of 10 periods) is also optimal for the other 19 trials—in short, how clearly feedback points in the direction of optimal behavior. Inconsistency in this case corresponds to a learning environment in which a good response strategy for one trial does not work well on other trials. Based on extensive research showing that learning depends heavily on how relevant performance feedback is to future behavior (e.g., Einhorn & Hogarth, 1978), we expect better learning in more *consistent* environments.

The second environmental factor we consider is *tolerance*, related to Hogarth's consequence of errors or "exactingness." Tolerance reflects the extent to which a wide range of strategies produce payoffs similar to that generated by the optimal strategy, or, in short, how strong the incentives are to correct suboptimal strategy. Perhaps surprisingly, previous research suggests that strong incentives can actually impede learning, especially for complex tasks. The dominant explanation is that the presence of incentives draws attention to how well a person is performing and away from a deeper understanding of the task (Hogarth, McKenzie, Gibbs, & Marquis, 1991). Although this can be beneficial for tasks requiring "simple, routine, unchanging responses and when circumstances favor the making of such responses quickly, frequently, and vigorously" (McCullers, 1978), it may become problematic when learning requires attention to a broad range of cues or experimentation (Easterbrook, 1959). For example, Ederer and Manso (2008) had participants make complex, multidimensional business decisions over many rounds and found that participants under a pay-for-performance incentive scheme showed *less* learning than those under fixed-payment schemes.

Given the complexity of our experimental task, we likewise expect more learning in more *tolerant* environments.

In sum, these two dimensions characterize how clearly an environment points an individual toward better decision-making, and the incentives it provides for moving in that direction. We hypothesize positive relationships between learning and both consistency and tolerance. Specifically, we expect that learning is most likely to occur when feedback is consistent and when the learning environment is tolerant.

In addition, by investigating the relationship between system variables (signal diagnosticity and transition probability) and learning variables (consistency and tolerance), we hope to shed light on when and why learning occurs. Therefore, we also hypothesize that the degree of learning will vary depending on how different combinations of system variables affect the consistency and tolerance of the environment.

It is important to note that other elements of a learning environment can affect its consistency and tolerance. For example, tolerance is also affected by the incentives scheme (e.g., payoff rules can be steeper or flatter—or completely flat as in fixed pay), and consistency is also affected by factors such as feedback frequency (e.g., providing feedback every three trials as in Lurie and Swaminathan (2009)) and clarity (e.g., providing only total payoffs but not regime information). This paper focuses only on the effects of the system variables, diagnosticity and transition probability, on the learning environment.

## LEARNING EXPERIMENT

### *Methods*

We recruited 72 University of Chicago participants, with the task advertised as a “probability estimation task” and assigned each to one of the six experimental conditions. The experiment was conducted using a specially designed Visual Basic program. The program began by introducing the statistical process used in the experiment, explaining the system variables ( $p_R$ ,  $p_B$ , and  $q$ ), how the computer would pick balls (i.e., the signals) from one of two bins (i.e., the regimes), and how the bin may switch. The program showed a schematic diagram of bin switching and then displayed four demonstration trials, each consisting of ten sequential random draws. For these demo trials only, participants saw the actual sequence of bins that generated each signal, and therefore if and when the process shifted from the red to the blue bin.

Participants were then told that after seeing each ball, their job was to estimate the probability that the system had switched to the blue bin (i.e., the probability that the regime had shifted) by entering any number between 0 to 100. The computer then gave a detailed explanation of the incentive procedure including payment curves as a function of estimated probability and whether the bin had actually switched to the blue bin yet. Participants completed two unpaid trials to better understand the incentive structure. After each trial, they were informed how much money they would have made or lost on that particular trial. Finally, participants completed the 20 trials for actual pay. At the end of each trial, participants received feedback about which bin generated each ball, earnings for that trial, and cumulative earnings in the experiment.

### *Results*

In this section, we summarize the basic results of our learning experiment. First we look at two measures of performance: (1) earnings, and (2) the mean absolute difference between empirical and Bayesian judgments. Next, to test the system-neglect hypothesis, we consider measures of reaction

or *changes* in probability judgments. Finally, we estimate a quasi-Bayesian model to provide a formal test of the system-neglect hypothesis. In all cases, we first consider how these measures vary across experimental conditions and then show how they change with experience over the course of the experiment.

*Earnings.* Recall that we paid participants on the difference between their subjective probability of change to the blue regime and the actual regime that period (1 if blue, 0 if red). The mean of this absolute difference was .24 (median = .07, sd = .32), generating mean variable earnings of \$10.97 total over 200 responses (range of \$-0.80 to \$14.48) and mean total earnings, including the fixed payment, of \$12.31 (range of \$4.20 to \$14.48). Table 1 presents the mean earnings in each condition and overall.

As a standard for comparison, a Bayesian agent would have earned \$14.20, including the fixed payment. Participants deviated from optimal earnings by \$1.89 (range of \$0.08 to \$10.34 across participants; see Table 1). A linear regression of the log deviation from optimal earnings on diagnosticity, transition probability, and their interaction found that participants in more precise ( $t = 3.42$ ,  $p < .001$ ) and stable ( $t = 1.82$ ,  $p = .07$ ) conditions earned closer to optimal, but found no interaction between the two parameters.

To investigate the possibility of improvement in earnings with experience, we first considered how earnings changed over the course of the 20 trials by dividing them into four quarters of 5 trials each. Table 1 lists the variable earnings by quarter, multiplied by four to be comparable to the overall variable earnings. There was some evidence of learning in all six conditions but learning was most dramatic between quarters 1 and 2, with performance improving at this stage in five of six conditions. Earnings also increased between quarters 2 and 3 in four conditions, and between quarters 3 and 4 in three conditions. The largest improvements occurred in the highly precise/stable condition ( $d = 9$  and  $q = .05$ ), with earnings per quarter increasing by \$0.44 from quarter 1 to 4. Impressively, earnings in the 4th quarter of this condition were only \$0.11 less than optimal. The only other condition that produced steady improvement was the highly precise/unstable condition ( $d = 9$ ,  $q = .20$ ).

We conducted a more rigorous test of learning by regressing the earnings by trial order. To do so, we took an observation to be a trial-condition pair so that each pair averaged the 120 (12 participants per condition  $\times$  10 periods per trial) earnings for that trial, again re-scaled to be comparable to overall earnings. Although the earnings by quarter indicated that learning is not linear or even necessarily monotonic, we assumed linearity here for simplicity. The regression coefficients are shown at the bottom of Table 1. Coefficients in five of the conditions were positive, though the improvement was significant only in the highly precise/stable condition. We also conducted the same analysis separately for each of the 72 participants. The percentages of participants who had positive regression coefficients are shown in Table 1. Coefficients were positive for 42 of the 72 participants overall ( $p = .16$ , sign test). Again, the only significant improvement was seen in the highly precise/stable condition, where 10 of the 12 participants had positive coefficients ( $p = .02$ , sign test).

*Mean Absolute Deviations.* Although the results on earnings show some evidence of learning, Bayesian judgments can produce lower earnings than non-Bayesian judgments for sequences of signals that are unrepresentative of the system that produces them. We therefore next considered the mean absolute deviation (MAD) between empirical and Bayesian judgments as a second, less noisy measure of performance that neither rewards nor punishes judgments based on luck. Table 2 shows results on MADs.

	Condition						Overall
	$d = 1.5$		$d = 3$		$d = 9$		
	$q = .05$	$q = .20$	$q = .05$	$q = .20$	$q = .05$	$q = .20$	
Empirical Earnings							
Variable Earnings	10.25	6.77	11.57	10.68	13.48	13.10	10.97
Standard Deviation	1.73	2.74	1.54	1.38	1.29	1.48	2.80
Fixed Payment	2.00	5.00	0.00	1.00	0.00	0.00	
Total Earnings	12.25	11.77	11.57	11.68	13.48	13.10	12.31
Bayesian Earnings							
Variable Earnings	12.44	9.55	13.12	13.14	14.58	14.36	12.86
Total Earnings	14.44	14.55	13.12	14.36	14.58	14.36	14.20
Bayesian - Empirical	2.19	2.78	1.56	2.46	1.10	1.26	1.89
Empirical Earnings by Quarter (scaled by 4)							
Quarter 1	9.74	6.54	11.67	10.21	12.41	12.71	10.55
Quarter 2	10.45	6.21	11.82	11.06	13.58	13.24	11.06
Quarter 3	11.26	7.41	11.39	11.41	13.77	13.09	11.39
Quarter 4	9.54	6.91	11.39	10.04	14.15	13.36	10.89
Difference between Bayesian and Empirical Earnings by Quarter (scaled by 4)							
Quarter 1	2.70	3.01	1.45	2.93	2.17	1.65	2.31
Quarter 2	1.99	3.34	1.30	2.08	1.00	1.12	1.80
Quarter 3	1.18	2.14	1.73	1.73	0.81	1.27	1.47
Quarter 4	2.90	2.64	1.73	3.10	0.43	1.00	1.97
Change in earnings by trial (scaled by 20)							
Regression Coefficient	.024	.036	-.004	.014	.098	.036	.034
Standard Error	.045	.039	.030	.030	.022	.025	.013
% Positive Learning	50%	58%	50%	67%	83%	50%	58%

Table 1: Mean Empirical and Bayesian earnings (variable and total) by condition. Also shown are standard deviations for empirical earnings, the difference between Bayesian earnings and empirical earnings (overall and by quarter), empirical earnings by quarter, linear regression coefficients of earnings change by trial, and percentage of participants with improving earnings. Data by quarter and trial are re-scaled for comparison.

	Condition						Overall
	$d = 1.5$		$d = 3$		$d = 9$		
	$q = .05$	$q = .20$	$q = .05$	$q = .20$	$q = .05$	$q = .20$	
Mean absolute deviations (MADs)							
Quarter 1	.189	.254	.138	.189	.123	.140	.172
Quarter 2	.191	.258	.135	.154	.085	.114	.156
Quarter 3	.186	.279	.128	.140	.075	.112	.153
Quarter 4	.185	.255	.131	.163	.056	.103	.149
Overall	.188	.261	.133	.161	.085	.117	.158
Change in MADs by trial							
Regression Coefficient	-.0007	.0004	-.0007	-.0019	-.0040	-.0024	-.0015
Standard Error	.0006	.0008	.0006	.0007	.0006	.0007	.0003
% Positive Learning	42%	33%	58%	75%	92%	83%	64%

Table 2: Mean absolute deviations (MADs) between empirical judgments and Bayesian judgments, by quarter and overall. Also, linear regression coefficients of MAD change by trial and percentage of participants with improving MADs.

Participants' responses showed the smallest MADs in the two highly precise conditions and the largest MADs in the two noisy conditions. Overall, MADs decreased over the four quarters, though the pattern was not uniform across conditions. The largest and most consistent improvements appeared again in the highly precise conditions, and in particular the highly precise/stable condition. There was also some learning in the moderately precise/unstable condition (though mostly from quarter 1 to 2), but virtually none in the remaining three conditions.

As before, we regressed the MADs by trial for each condition (see Table 2). The coefficients were negative (suggesting learning) for five of the conditions and significant in the two highly precise conditions and the moderately precise/unstable condition. The strongest learning again appeared in the highly precise/stable condition. Separate regressions for each participant showed a similar but stronger pattern to the analysis in Table 1, with 46 of the 72 participants having negative coefficients ( $p < .02$ , sign test) and 11 of 12 participants improving in the highly precise/stable condition ( $p < .01$ , sign test). Otherwise, 10 of 12 participants improved in the highly precise/unstable condition ( $p = .02$ ) and 9 of 12 participants improved in the moderately precise/unstable condition ( $p = .08$ ). Overall, we found somewhat stronger evidence of learning for this more precise measure of performance.

*Measures of reaction.* Whereas our analyses of earnings and MADs demonstrated learning differences across conditions, the system-neglect hypothesis specifies how empirical probability judgments *react* to indications of change (i.e., blue signals) rather than their absolute levels. To test for system neglect, we therefore compared participants' reactions (i.e., the change in probability

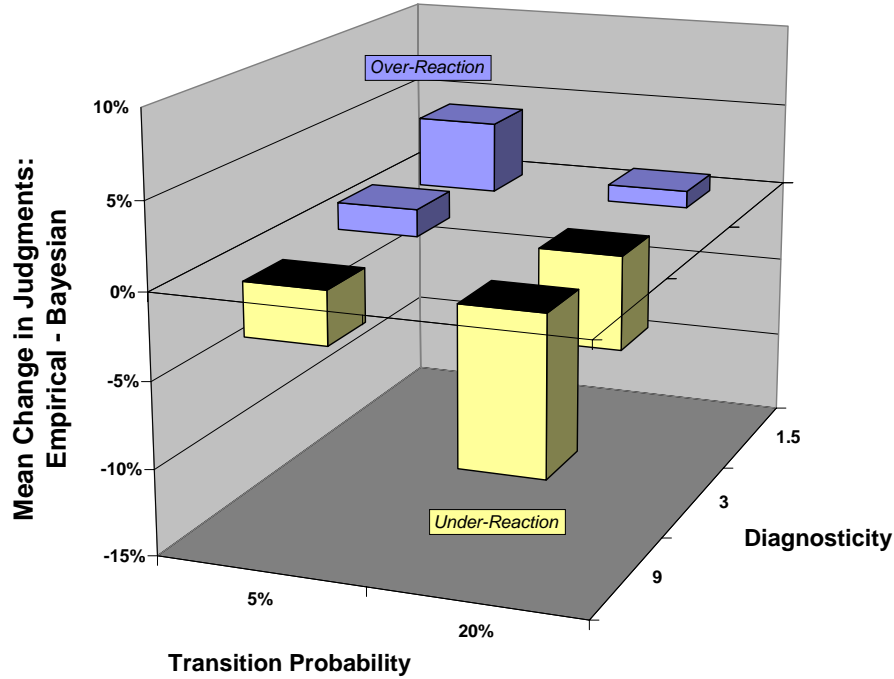


Figure 1. Over- and underreaction to blue signals, by condition, as measured by the mean difference between the change in empirical probability judgments and Bayesian reaction.

judgments from period  $i - 1$  to period  $i$ ) to the Bayesian reaction assuming that the previous period's probability judgment is the correct "prior" (see MW for more details). Importantly, this approach focuses only on *reactions*, granting participants their priors, whether accurate or not, and evaluating only how their judgments react to new information. We term errors in reaction as the difference between Bayesian and empirical reactions, with *underreaction* indicating an empirical reaction that is less positive than the Bayesian reaction, and *overreaction* indicating the opposite. Recall that the system-neglect hypothesis predicts a greater tendency to underreact in more precise, less stable conditions, and to overreact in noisier, stabler conditions.

Figure 1 depicts the mean error in reactions to blue signals by condition (red signals are indicators of "non-change" and exhibit a different gradient; see Massey & Wu, 2005a, p.934-935). As predicted by the system-neglect hypothesis, and replicating MW, the greatest underreaction occurred in the southeast-most cell ( $d = 9$ ,  $q = .20$ ), while the greatest overreaction occurred in the northwest-most cell ( $d = 1.5$ ,  $q = .05$ ). For all 12 pairwise comparisons between conditions, underreaction increased monotonically with diagnostics and transition probability.

Figure 2 plots the same errors in reactions to blue signals by quarter. Note that the degree of system neglect is most pronounced in the first quarter but remains significant in the remaining three quarters. Consistent with the analyses in the previous sections, it appears that the learning that does take place occurs mostly in the highly precise conditions, and from quarter 1 to quarter 2.

*Quasi-Bayesian Estimation.* We next follow MW in further analyzing the pattern of reactions in Figure 1 by estimating a *quasi-Bayesian* model to test for learning to detect change. This analysis allows us to formally rule out the possibility that the hypothesized pattern could be an artifact of

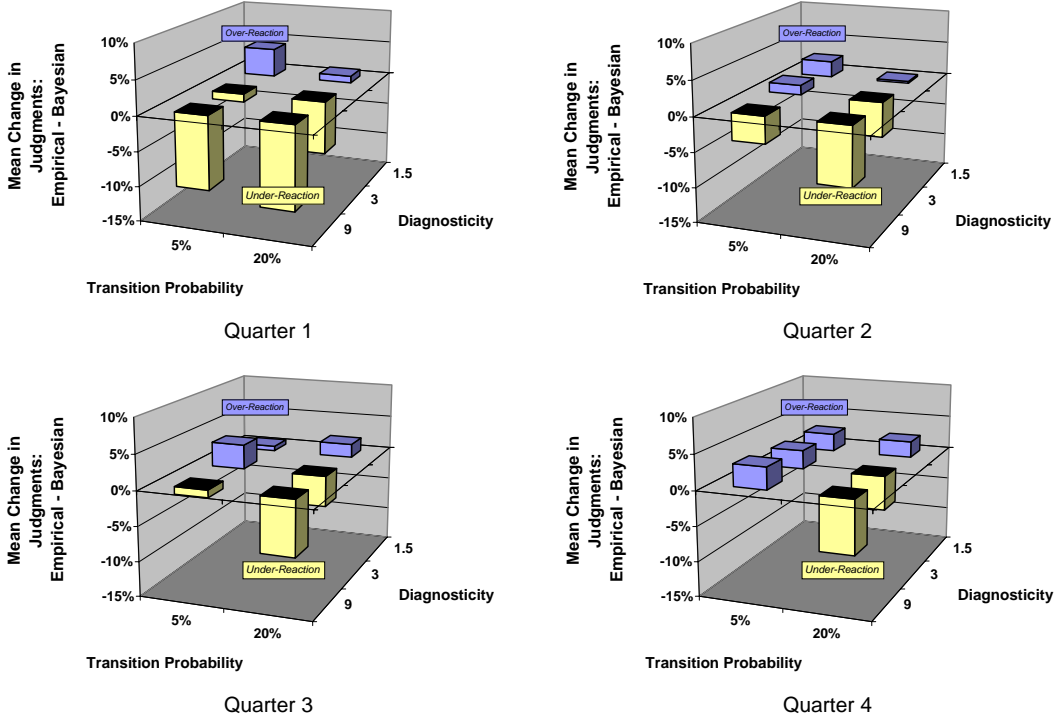


Figure 2. Over- and underreaction, by condition and quarter, as measured in Figure 1.

the specific sequences of signals. Examining learning through the lens of the quasi-Bayesian model also allows us to say more about *what* participants are learning.

The quasi-Bayesian model is a generalization of Equation (1) that explicitly allows for non-optimal sensitivity to transition probability and diagnosticity. We do so by adding two parameters,  $\alpha$  (sensitivity to transition probability) and  $\beta$  (sensitivity to diagnosticity):

$$\frac{p_i^e}{1 - p_i^e} = \left( \frac{1 - (1 - \alpha_m q)^i}{(1 - \alpha_m q)^i} \right) \sum_{j=1}^i \frac{q(1 - q)^{j-1}}{1 - (1 - q)^j} d^{\beta_n} \left[ i+1-j - \left( 2 \sum_{k=j}^i b_k \right) \right] \quad (2)$$

Note that  $\alpha_m < 1$  and  $\beta_n < 1$  reflect insensitivities to transition probability and signal diagnosticity, respectively, whereas  $\alpha_m = \beta_n = 1$  returns the Bayesian expression in Equation (1). The subscripts  $m$  and  $n$  denote the fact that  $\alpha$  and  $\beta$  must vary by condition to produce system neglect: in fact, it requires the following monotonic ordering of these parameters:  $\alpha_{.05} > \alpha_{.20}$  and  $\beta_{1.5} > \beta_3 > \beta_9$  (see MW for a more complete discussion of this model).

We first ran the model by pooling all the data across individuals and conditions and estimating a single omnibus nonlinear regression, thereby modeling the behavior of a representative participant (McFadden, 1981). Results for the means of individual estimates were qualitatively similar. Figure 3 shows the estimates of  $\alpha_m$  and  $\beta_n$  as well as parameters for complete system neglect. Here, lower levels imply greater conservatism. The estimated parameters were ordered as predicted by system neglect, with all pairwise differences significant at  $p < .001$ . In addition, there was a mixed pattern of conservatism and radicalism in both parameters, consistent with the pattern of over- and underreaction depicted in Figure 1. Importantly, whereas the slopes were much steeper than

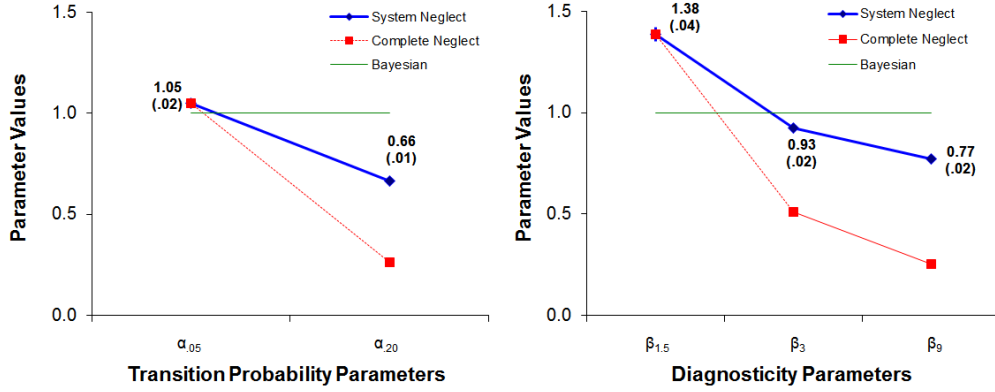


Figure 3. System neglect parameter estimates by condition

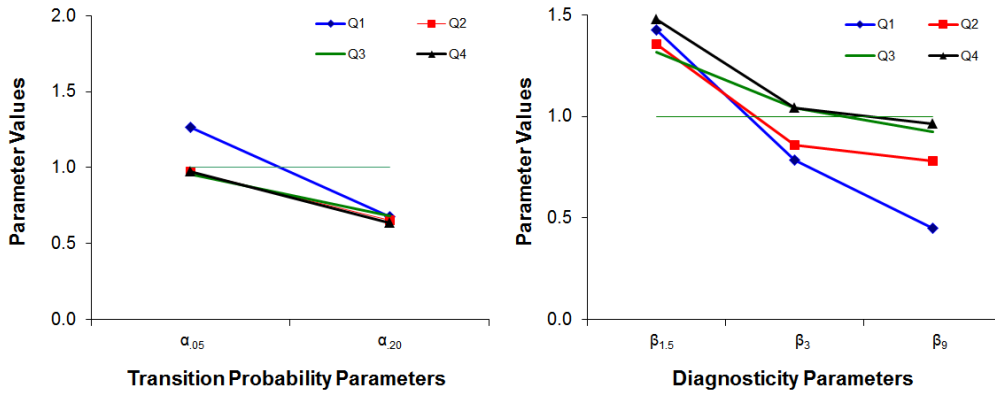


Figure 4. System neglect parameter estimates by quarter and condition

Bayesian, they were less shallow than the slope of the complete-neglect curves. In other words, although participants were not sufficiently sensitive to system variables, they were not completely insensitive, either.

To investigate whether learning was responsible for the observed differences in system neglect, we estimated  $\alpha_m$  and  $\beta_n$  as before, adding dummies for quarter. Figure 4 shows these estimates for all four quarters. Several inferences can be drawn from this figure. First, learning occurred for stable ( $q = .05$ ) conditions but not unstable ones ( $q = .20$ ). Second, the parameter estimates for both moderately ( $d = 3$ ) and highly precise ( $d = 9$ ) conditions appeared to converge toward the Bayesian standard. This convergence was particularly dramatic for the highly precise conditions. In contrast, there was virtually no learning in the noisy conditions ( $d = 1.5$ ). Third, most of the learning happened early, between quarters 1 and 2. These results are again highly consistent with the earlier analyses on earnings and MADs.

*Summary.* Whether analyzing earnings, mean absolute deviations from Bayesian estimates and reactions, or estimates of the quasi-Bayesian model, the results in this section were all consistent with two general patterns: (i) learning tended to occur early, if at all; and (ii) learning was most prevalent in the highly precise/stable condition ( $d = 9, q = .05$ ), followed by the highly

precise/unstable condition ( $d = 9, q = .20$ ). There was modest learning in the moderately precise/unstable condition ( $d = 3, q = .20$ ) and hardly any in the other three conditions.

Although our findings also replicated the predicted pattern of over- and underreaction, system neglect is silent on why learning varies in this way across the experimental conditions. In order to better understand this variation, we next turn to a more direct analysis of learning environment characteristics.

## LEARNING ENVIRONMENT

We have analyzed learning thus far in terms of our system variables, diagnosticity and transition probability. However, it might be more informative to examine the learning environments faced by our participants more directly. In this section we investigate two environmental factors that critically affect how we learn: the quality of feedback (consistency) and the consequence of errors in judgment (tolerance). We first examine how these factors vary as a function of the system variables, and then explore their impact on learning. Our analysis shows why certain environments are more conducive to learning than others.

Note that we have chosen to investigate the environment explicitly, rather than the learning process itself. This is distinct from other research on learning in stochastic environments that directly examines a reinforcement learning process (Bush & Mosteller, 1953; Camerer & Ho, 1999; Cross, 1983; Erev & Roth, 1998). We adopt our modeling approach for two reasons. First, just as we focus on system variables to understand when system neglect leads to over- and underreaction, we focus on environmental factors to understand what leads to changes in system neglect over time. Second, the judgments we study are dynamic. As such, judgments are meaningful only in context. For example, a response of .50 in the first period of a trial is very different from a response of .50 after three blue signals. Given the unwieldy number of possible scenarios we would have to model learning for, the most obvious way to apply reinforcement learning in this setting would be at the parameter level. That is, rather than payoffs reinforcing certain responses in various scenarios, learning would move  $\alpha$  and  $\beta$  toward more optimal parameter values. While such an approach seems reasonable, it would be quite a departure from previous research. We return to this point below.

### *Operationalization*

First, we consider the quality of feedback. Whereas feedback quality may include many features such as clarity, relevance, and timeliness, these are constant across our experimental conditions. The key feature that varies is the *consistency* of feedback, or how well one trial's optimal parameters perform for other trials in that condition. Feedback is consistent when performance on subsequent trials reinforces the optimal parameters deduced in earlier trials.

To operationalize feedback consistency, let  $e_t(\alpha_t, \beta_t)$  denote the earnings in trial  $t$  from responding in a quasi-Bayesian fashion with  $\alpha_t$  and  $\beta_t$  (see Equation (2)) on that trial. Also, let  $\alpha_t^*$  and  $\beta_t^*$  be the optimal parameters maximizing earnings in trial  $t$ . For trials without a unique optima (e.g.,  $\beta_t^*$  is not uniquely determined in trials with no red signals and neither  $\beta_t^*$  nor  $\alpha_t^*$  are determined in trials with no blue signals), we define  $\alpha_t^*$  and  $\beta_t^*$  as the midpoint of the optimal plateau. Under this definition, feedback is more *consistent* for large values of Equation (3):

$$Consistency = \frac{1}{19 \times 20} \sum_t \sum_{t' \neq t} \frac{e_{t'}(\alpha_t^*, \beta_t^*)}{e_{t'}(\alpha_{t'}^*, \beta_{t'}^*)} \quad (3)$$

	Condition					
	$d = 1.5$		$d = 3$		$d = 9$	
	$q = .05$	$q = .20$	$q = .05$	$q = .20$	$q = .05$	$q = .20$
Consistency (Mean performance using optimal parameters of other series)	61.5%	18.0%	53.1%	34.2%	66.8%	73.8%
Tolerance (Percent of parameter space within $\tau$ of optimal)						
$\tau = \$0.50$	0.4%	1.3%	1.6%	3.1%	8.9%	7.9%
$\tau = \$1.00$	0.6%	3.0%	2.9%	9.8%	32.8%	31.1%
$\tau = \$1.50$	1.2%	5.4%	4.0%	41.3%	50.0%	44.0%
$\tau = \$2.00$	1.4%	8.9%	5.1%	63.6%	54.1%	57.4%

Table 3: Analysis of learning environments. Consistency indicates the mean performance of each trial’s optimal parameters on all other trials in the condition. Higher values indicate more consistent feedback. Tolerance indicates what percentage of the  $0 \leq \alpha \leq 5$  and  $0 \leq \beta \leq 5$  space produces total earnings within a threshold,  $\tau$ , of the optimal parameters for that condition,  $\alpha^{**}$  and  $\beta^{**}$ . Higher values indicate environments more tolerant of errors in judgment.

This measure reflects the mean degree to which the optimal quasi-Bayesian parameters from each trial are reinforced in the other trials, relative to the optimal earnings achievable on those trials.

Second, we consider the consequences of inaccurate judgments. We define an environment as more *tolerant* if errors are not punished very harshly relative to correct judgments. We operationalized tolerance as the degree to which incorrect  $\alpha$  and  $\beta$  parameters generate earnings within a certain threshold of the maximum earnings. In other words, more tolerant environments have flatter maxima. Let  $E(\alpha, \beta)$  denote the earnings from using parameters  $\alpha$  and  $\beta$  for all 20 trials in a particular condition, and let  $\alpha^{**}$  and  $\beta^{**}$  denote the parameters that maximize total earnings for that condition. For each condition, we calculated  $E(\alpha, \beta)$  for all combinations of  $0 \leq \alpha \leq 5$  and  $0 \leq \beta \leq 5$ , thus generating the earnings space in  $\alpha$ - $\beta$  dimensions. We chose to cap the upper limit for each variable at 5 due to diminishing marginal effects of higher values on judgment beyond this point. Other reasonable upper limits returned similar results. We then determined what percentage of combinations of  $\alpha$  and  $\beta$  produce earnings within some threshold  $\tau$  of the maximum earnings, as specified in Equation (4):

$$E(\alpha^{**}, \beta^{**}) - E(\alpha, \beta) \leq \tau. \quad (4)$$

#### *System Variables and Learning Environment Factors*

Table 3 shows consistency measures, as well as tolerance measures for thresholds ( $\tau$ ) ranging from \$0.50 to \$2.00. Both consistency and tolerance vary systematically with the system variables. Feedback is most consistent in the highly precise conditions, followed by the noisy/stable condition, and least consistent in the noisy/unstable condition. A linear regression of consistency on diagnosticity, transition probability, and their interaction revealed that stable systems are more consistent than unstable ones ( $p < .05$ ), but that this effect is attenuated for moderately precise conditions

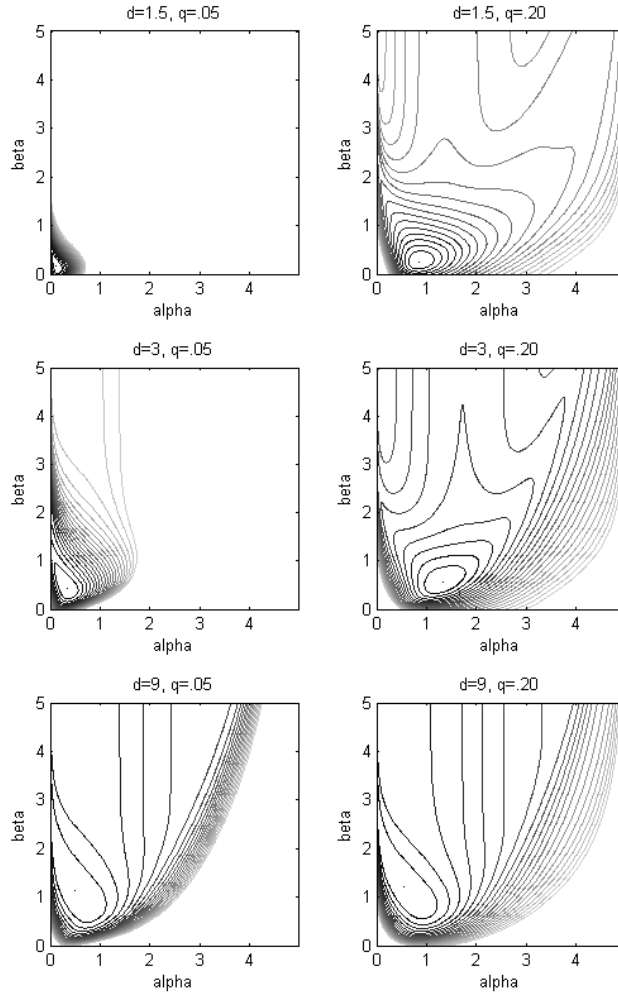
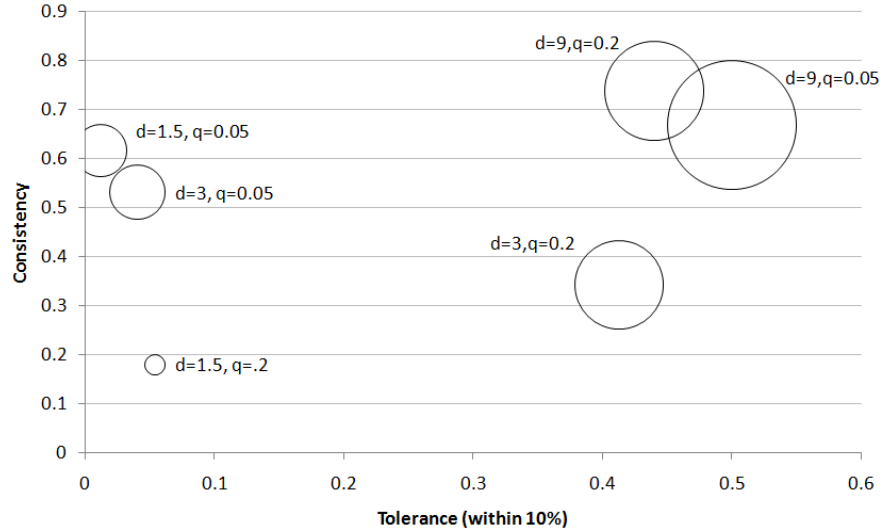


Figure 5. Earnings space contours maps in  $\alpha$ - $\beta$  dimensions assuming quasi-Bayesian model. For each condition, the plots show the total earnings across the 20 trials for all combinations of  $0 \leq \alpha \leq 5$  and  $0 \leq \beta \leq 5$ . Contours are spaced at 25¢ intervals from maximum earnings.

and slightly reversed for highly precise ones ( $p < .05$ ).

To better visualize the measures of tolerance, Figure 5 depicts tolerance as contour maps of the earnings space. These contour maps show that the earnings space is remarkably flat for the highly precise conditions, and relatively flat for the moderately precise/unstable condition. That is, these conditions have fairly flat maxima with large regions of  $\alpha$  and  $\beta$  producing near-optimal earnings. On the other hand, the earnings space is quite steep for the other two stable conditions ( $d = 1.5$ ,  $q = .05$  and  $d = 3$ ,  $q = .05$ ). Whereas half of the earnings space in the highly precise/stable condition perform within \$1.50 (roughly 90%) of the optimal earnings, only 5% of earnings space for the other two stable conditions perform within this threshold. That is, the noisier conditions had sharper maxima. A regression of tolerance on diagnosticity, transition probability, and their interaction revealed only a positive main effect of diagnosticity, regardless of  $\tau$  level ( $p < .05$  for all, and  $p < .01$  for  $\tau \leq \$1.00$ ). Systems with more precise signals produce more tolerant learning environments.



*Figure 6.* Learning across different conditions as a function of consistency and tolerance. The diameter of each circle is proportional to the amount of learning as indicated by the reverse scored regression coefficient of MAD learning. Consistency is measured by the mean performance using each trial’s optimal parameters on all other trials and tolerance is measured by the percentage of  $0 \leq \alpha \leq 5$  and  $0 \leq \beta \leq 5$  that perform within \$1.50 (roughly 90%) of the optimal earnings.

In general we see that the system variables influence these environmental factors in systematic but complicated ways. It is precisely because of this complexity that consistency and tolerance may relate more strongly to learning than do the system variables. More importantly, quantifying consistency and tolerance allows us to draw theoretical relationships between learning environments and learning where the system-neglect hypothesis does not.

#### *Environmental factors and learning*

Recall that we predicted that both a learning environment’s consistency and tolerance would contribute to better learning for complex tasks such as change detection. We therefore related consistency and tolerance to the empirical learning we observed across conditions using the change in mean absolute deviation over time (MAD learning, for short) measure as reported in Table 2. Other measures of learning—earnings, deviations from Bayesian reactions, and estimates of the quasi-Bayesian model—produced similar results.

Figure 6 plots MAD learning as a function of consistency and tolerance, using  $\tau = \$1.50$  (row 6 in Table 3), or the area within roughly 90% of the optimal earnings. We reverse scored MAD learning so that positive coefficients and larger diameter circles correspond to more empirical learning. The figure reveals a strong relationship between learning and tolerance: In 13 of the 15 pairwise comparisons, learning was higher for more tolerant environments. There was also a positive relationship between learning and consistency, though a somewhat weaker one: In 11 of the 15 pairwise comparisons across conditions, learning was higher for more consistent environments.

We formally examined these relationships by regressing the reversed MAD learning coefficients on consistency and tolerance. Consistent with the graphical analysis, tolerance had a significant positive effect on learning ( $t = 4.75, p = .01$ ), whereas consistency had only a marginally positive effect ( $t = 2.59, p = .08$ ). To assess the robustness of these effects, we replicated this regression

using other thresholds of tolerance at intervals of  $\tau = \$0.10$  up to  $\$3.00$ . In these 30 regressions, the effects of tolerance and consistency were always in the same direction, though the results for consistency dropped below standard levels of significance for  $\tau < \$1.40$ .

A positive relationship between consistency and learning is consistent with intuition, but the relative weakness of the relationship is surprising. Although we can only speculate about why we found a marginally significant relationship (e.g., the limited sample size and measurement error in both the independent and dependent variables surely play some role), perhaps the most likely explanation is that our environments do not span a broad enough range of consistencies. It seems that highly consistent feedback—beyond that which we evaluated—would enable learning in even the least tolerant environments.

The positive relationship between tolerance and learning, on the other hand, is less intuitive. Our experimental paradigm examines the relation between learning and incentives by separating the feedback participants receive about their performance from the incentive they have to change it. In the language of the contour maps presented in Figure 5, feedback points in the direction of the peak, while incentives indicate the potential change in elevation achieved by moving in that direction. The negative relationship between incentives (i.e., less tolerance) and learning indicates that, holding feedback constant, individuals learned better with flatter maxima. That is, they seemed to improve more with experience when they had less to gain from doing so!

This possibility was discussed by Hogarth et al. (1991). Building on research on how incentives affect learning in complex environments, they argued that as the consequences of decisions increase, attention is diverted “from inference to evaluation” (Hogarth et al., 1991, p. 735). That is, when consequences are high in a complex environment, individuals focus more on rewards and punishments, rather than how to best perform the task. Easterbrook (1959) suggests that one reason this may impair learning is that the high drive state induced by strong incentives restricts attention to a limited range of cues. Yet, a broad search is often exactly what is required for learning in a complex environment, as we saw in Figure 5. While we are cautious in generalizing too much from this observation, it is provocative how strongly it runs counter to economic intuition.

In sum, this analysis shows that experience alone is not sufficient for learning. Rather, the impact of experience depends on the learning environment. Drawing on research in psychology, we found that environments vary significantly in their conduciveness to learning. The pattern is stark, despite a small sample and sparse stimuli: Learning depends on the consistency of feedback and even more so on the environment’s tolerance to errors.

## AN ALTERNATIVE CHARACTERIZATION OF LEARNING ENVIRONMENTS

One shortcoming of the preceding analysis is that characterizing each learning environment in terms of tolerance and consistency as defined in quasi-Bayesian  $\alpha$ - $\beta$  space may be disconnected from participants’ actual behavior. The quasi-Bayesian model only acts as an “as-if” model, and was designed to provide a formal test of the system-neglect hypothesis. That is, we do not expect any participant to actually calculate Equation (2). In this section, we discuss an alternative analytic strategy for understanding learning by viewing behavior through the lens of a computationally simpler and thus more psychologically plausible strategy, which we term the *delta-epsilon heuristic*.

In brief, this purely linear heuristic calls for a fixed additive response to each signal of change (i.e., a blue signal), and a different fixed response to each signal of non-change (i.e., a red signal). More specifically, the heuristic starts by returning a default or baseline ( $\gamma$ ) probability that a regime

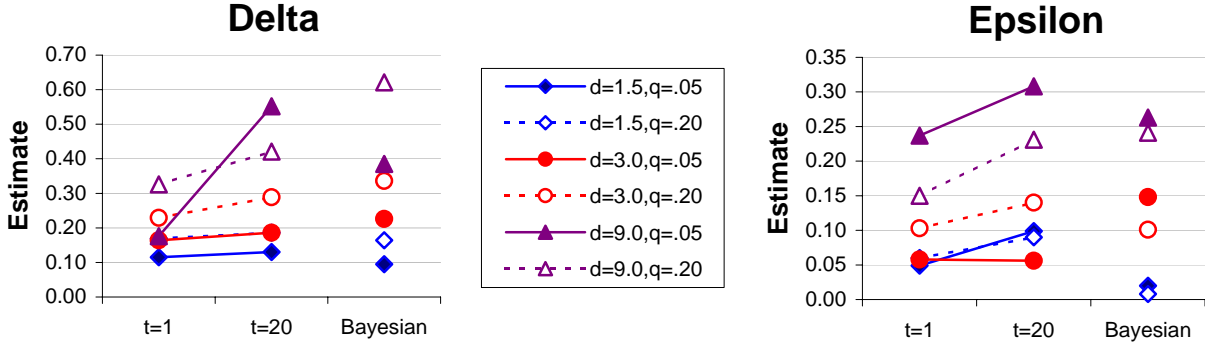


Figure 7. Bayesian and empirical estimates for  $\delta$  and  $\varepsilon$  across conditions, estimated for  $t = 1$  and  $t = 20$ , using the linear learning representative agent model.

change has occurred. Once the first blue signal is observed, it adds  $\delta$  to that probability for each blue signal, and subtracts  $\varepsilon$  for each red signal, with the probability constrained between 0 and 1. Therefore,  $\delta$  is the response to an indication of change, whereas  $\varepsilon$  is the response to an indication of non-change. Mathematically, the delta-epsilon heuristic is defined by Equation (5):

$$p_t = \begin{cases} \gamma & \text{if } \sum_{1}^t b_t = 0, \\ \min(1, p_{t-1} + \delta) & \text{if } \sum_{1}^t b_t > 0 \text{ and } b_t = 0, \\ \max(0, p_{t-1} - \varepsilon) & \text{if } \sum_{1}^t b_t > 0 \text{ and } b_t = 1. \end{cases} \quad (5)$$

We analyze behavior through the lens of this heuristic for a number of reasons. Most importantly, whereas the quasi-Bayesian model is far too complex to represent participants’ strategies, the delta-epsilon heuristic is sufficiently simple and intuitive to plausibly approximate the actual reasoning strategies used to arrive at empirical outcomes. We believe this approach can provide deeper insights into individual behavior, and is complementary to the more formal approach of the quasi-Bayesian model. There is a fruitful tradition of studying simple heuristics like this in a variety of decision settings, including game theory (Costa-Gomes, Crawford, & Broseta, 2001), judgment under uncertainty (Anderson, 1965; Gigerenzer & Goldstein, 1996; Hogarth & Karelaia, 2007; Shanteau, 1975), and multi-attribute choice (Gigerenzer, Todd, & The ABC Group, 1999; Hogarth & Karelaia, 2005; Payne, Bettman, & Johnson, 1993). Moreover, these heuristics can approximate normative behavior (Dawes, 1979; Dawes & Corrigan, 1974). Indeed, the delta-epsilon heuristic can closely approximate Bayesian updating, yet is also able to fit individual and aggregate level probability judgments fairly well. We present evidence on these points below.

#### Delta-Epsilon Heuristic Fits and Learning

We first consider how well the simple delta-epsilon heuristic tracks Bayesian updating. For each condition, we used nonlinear regressions to fit  $\delta$ ,  $\varepsilon$ , and  $\gamma$  to Bayesian judgments in each period. Figure 7 shows Bayesian parameter estimates for  $\delta_b$  and  $\varepsilon_b$  (values of  $\gamma_b$  do not vary much). Despite its computational simplicity, the delta-epsilon heuristic was able to closely approximate Bayesian

judgments. The mean absolute deviation between Bayesian judgments and judgments produced by the delta-epsilon heuristic using the Bayesian parameter estimates ranged from .018 to .069, with high correlations between probabilities based on the delta-epsilon heuristic and Bayesian updating ( $r = .948$  to  $.995$ ). For all conditions, these estimates produced earnings that were nearly identical to Bayesian, with absolute deviations in earnings ranging from \$0.01 to \$0.31. In addition,  $\delta_b$  and  $\varepsilon_b$  varied dramatically across conditions, indicating that the degree of reaction should be sensitive to both the transition probability and diagnosticity. For example,  $\delta_b$  is six times higher in the highly precise/unstable system than it is in the noisy/stable one (.621 vs. .095). Overall,  $\delta_b$  and  $\varepsilon_b$  was higher for more precise signals, whereas  $\delta_b$  was higher and  $\varepsilon_b$  was lower with higher instability.

We next fit the delta-epsilon heuristic to the empirical judgments at the aggregate level. Individual fits for the 72 participants yielded similar results. Recall that Bayesian updating requires that reactions to a blue signal differ across conditions. Indeed, Bayesian  $\delta_b$  ranged from .095 to .620, and  $\varepsilon_b$  ranged from .008 to .263. In contrast, and consistent with system neglect, aggregate empirical estimates  $\delta_e$  and  $\varepsilon_e$  were considerably more compressed:  $\delta_e$  ranged from .122 to .380 and  $\varepsilon_e$  ranged from .057 to .277. In terms of the delta-epsilon model, underreaction occurs when  $\delta_b > \delta_e$ , and overreaction if the opposite holds. As predicted by system neglect, we found the most underreaction in the highly precise/unstable condition and the most overreaction in the noisy/stable condition.

To examine the possibility of learning under this heuristic, we next estimated a linear learning model, in which  $\delta_e(t)$ ,  $\varepsilon_e(t)$ , and  $\gamma_e(t)$  were allowed to change linearly across trials. Learning occurs if  $\delta_e(t) \rightarrow \delta_b$ ,  $\varepsilon_e(t) \rightarrow \varepsilon_b$ , or  $\gamma_e(t) \rightarrow \gamma_b$  as  $t$  increases. Results of these regressions showed that strong system neglect at  $t = 1$  tended to dissipate but not disappear over time. Figure 7 plots the evolution of the delta-epsilon parameters with experience, showing that while parameter values were too compressed in early trials, they moved in the direction of the Bayesian values over time. However, this learning pattern was more evident for  $\delta_e$  than for  $\varepsilon_e$ , consistent with the intuition that people learn to react to more salient signals of change than signals of non-change.

Individual linear learning model estimates fit empirical estimates fairly well—the mean absolute deviation between the heuristic and empirical judgments was only .15—and revealed a similar pattern of learning. Whereas 83% of participants moved in the direction of  $\delta_b$  and  $\varepsilon_b$  in the highly precise/stable condition and 92% in the highly precise/unstable condition, only 67% did so in the noisy/unstable condition.

Finally, we characterized the consistency and tolerance of each learning environment in  $\delta$ - $\varepsilon$  dimensions as we did for the quasi-Bayesian model and related these alternate definitions of consistency and tolerance to MAD learning. Results were nearly identical to those from the quasi-Bayesian analyses. In 12 of 15 pairwise comparisons, learning was higher with more consistent feedback, and in 14 of the 15 pairwise comparisons, learning was higher when environments were more tolerant. Both consistency and tolerance (again with varying levels of significance on the consistency coefficient depending on choice of  $\tau$ ) had significant positive effects on learning ( $t = 2.92$  and  $t = 4.31$ , respectively,  $p < .01$ ).

Therefore, the psychologically plausible and computationally reasonable delta-epsilon heuristic produced essentially identical results to the analyses on the quasi-Bayesian model. In addition to verifying the robustness of our learning environment analysis, the delta-epsilon heuristic suggests a potential mechanism for learning by adjusting the heuristic’s parameters. For example, one can imagine providing participants with the form of the delta-epsilon heuristic without providing the optimal parameter values. Then, learning from experience can consist of actually adjusting  $\delta$  and  $\varepsilon$  towards higher earning values. By providing a framework for participants to make simple,

methodical calculations and to structure feedback in more easily processable manner, the delta-epsilon may provide an invaluable tool for learning to detect change. Future research is necessary to explore this possibility.

## GENERAL DISCUSSION

This research was motivated primarily by the desire to pit system neglect against learning. Given sufficient experience with a single system, would participants grow more sensitive to its key features and learn to avoid the over- and underreaction driven by system neglect? The answer, while depending on the learning environment, is not all that positive. Across a wide range of environments, participants displayed the same pattern observed in Massey and Wu (2005a): relatively more overreaction in noisy and stable systems, and relatively more underreaction in precise and unstable systems. Our participants gained experience sufficient for some learning in some conditions. However, where we observed learning, it was not enough to overcome system neglect, and in some conditions we observed no learning at all.

This observation underscores both the robustness of system neglect and the challenge of learning. Experience alone is often not sufficient to overcome judgmental biases. Rather, the impact of experience is moderated by one's environment. In our experiment, we saw suggestions of learning across most conditions, but reliable evidence in only three of the six conditions. Learning to avoid over- and underreaction depended to some extent on the consistency of feedback and strongly on the tolerance of the environment to errors. Whereas we saw modest improvements in the early trials of most of the conditions, this improvement persisted only in highly tolerant environments or in moderately tolerant environments with highly consistent feedback. That is, consistent with Hogarth et al. (1991), we found that stronger incentives for improved performance actually impair learning.

This paper makes three main contributions. First, we extended the literature on change-point detection by establishing the robustness of the system-neglect hypothesis, along with its implied pattern of over- and underreaction. Second, we developed and tested explanations of learning in dynamic settings. In so doing, we demonstrated the critical role of the learning environment in moderating the impact of experience on performance. Third, we proposed that the complex Bayesian calculations of change detection can be simplified using a simple linear heuristic, showed that this heuristic does a good job of fitting empirical patterns, and suggested that it can help people better learn from experience by helping them process feedback.

### *Learning in stochastic environments*

Our characterization of learning environments in terms of tolerance to suboptimal behavior and consistency of feedback is generalizable to a wide range of stochastic environments. For example, a classic decision task in operations is the newsvendor's problem of deciding how many newspapers to stock in the face of uncertain demand (Arrow, Harris, & Marschak, 1951). In this problem, the retailer must balance the cost of overstocking (the unit cost of the item minus salvage costs) and the cost of understocking (the lost profit and possible ill will from irate customers). This problem has an elegant and relatively simple solution: stocking the "critical fractile" that balances the cost of understocking and overstocking. Nonetheless, empirical studies have shown relatively poor performance and modest learning at best (Schweitzer & Cachon, 2000; Bolton & Katok, 2008).

There are straightforward ways to operationalize the learning environment variables in the newsvendor problem. For example, given a fixed optimum at the critical fractile, higher variance

in the demand distribution leads to lower consistency of feedback. Tolerance is given by the size of the near-optimal earnings peak, which is also related to the location of the critical fractile. For uniform demand distributions, higher critical fractiles correspond to more tolerant environments.

Similar analysis can be performed for essentially any stochastic decision task, such as the Monty Hall problem and Multi-Armed bandit problems (Friedman, 1998; Sutton & Barto, 1998). Operationalizing these task environments in terms of tolerance and consistency can generate novel predictions for how varying the task will affect learning. For example, Ederer and Manso (2008) ran a laboratory experiment using a more complex variant of a three armed-bandit problem where participants selected one of three locations to sell lemonade as well as four features of the lemonade. Consistent with our predictions for tolerance, they found that pay-for-performance was detrimental to performance in finding the global optimum because it caused participants to explore fewer locations. While the authors provide little detail about the specific payoff functions in the game, variations such as flatter maximum payoffs across locations (tolerance) and simpler relationships between product features (consistency) would provide additional mechanisms for improved learning.

### *Organizational decision-making*

The influence of environmental factors on learning has direct implications for organizations. Obviously it would be helpful to improve consistency by increasing the quality or quantity of feedback available to a decision maker. Unfortunately firms often do not or cannot control the feedback available in their environment. However, they are able to improve decision-makers' *attention* to feedback. This can happen via enhanced record-keeping or through activities explicitly aimed at learning from the past (Cyert & March, 1963). Another approach is the use of policies to restrict decision-makers' freedom (Heath, Larrick, & Klayman, 1998) with the goal of avoiding "noise chasing" (i.e., overreacting to inconsistent feedback). Both of these approaches—learning programs and policy-based decisions—are ways to improve institutional memory, an adaptive response to inconsistent environments.

The role of tolerance also has organizational implications. Firms exert considerable discretion over tolerance through their reward and punishment policies. Strong incentives, especially narrow or short-term incentives, decrease tolerance by creating an environment that is very responsive to success and failure (Baron & Hershey, 1988; Bukszar & Connolly, 1988). These incentives take many forms, from promotion and pay-for-performance to informal attitudes toward failure. Improving tolerance does not mean firms should avoid incentives altogether. Rather, they should be judicious in their use, being especially attentive to periods and situations when learning is critical. Explicitly constructing periods in which "performance" is not punished or rewarded (e.g., Ederer & Manso, 2008), for example, can be helpful. Another approach is to reward learning explicitly, independently from performance. A firm must worry about complacency, of course, but our results underscore the tradeoff between strength of incentives and learning.

## CONCLUSIONS

Our paper establishes the robustness of system neglect in change-point detection and demonstrates the relationship between environmental conditions and learning. In the end, we are somewhat sober about the ability of individuals to avoid systematic over- and underreaction in non-stationary environments. However, we are also encouraged by the *possibility* of learning. Together these sentiments suggest that one of the most important directions for future research is to understand how different decision environments impact the potential for learning.

## References

- Anderson, N. H. (1965). Averaging versus adding as a stimulus-combination rule in impression-formation. *J. Experiment. Psych.*, *70*(4), 394-400.
- Arrow, K. J., Harris, T., & Marschak, J. (1951). Optimal inventory policy. *Econometrica*, *19*(3), 250-272.
- Ball, L. (1995). Time-consistent policy and persistent changes in inflation. *J. Monetary Econom.*, *36*(2), 329-350.
- Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *J. Financial Econom.*, *49*(3), 307-343.
- Baron, J., & Hershey, J. (1988). Outcome bias in decision evaluation. *J. Personality Soc. Psych.*, *54*(4), 569-579.
- Barry, D. M., & Pitz, G. F. (1979). Detection of change in nonstationary, random sequences. *Organizational Behavior and Human Performance*, *24*(1), 111-125.
- Blinder, A. S., & Morgan, J. (2005). Are two heads better than one? Monetary policy by committee. *J. Money, Credit, and Banking*, *37*(5), 789-811.
- Bolton, G. E., & Katok, E. (2008). Learning by doing in the newsvendor problem: A laboratory investigation of the role of experience and feedback. *Manufacturing Service Oper. Management*, *10*(3), 519-538.
- Bowler, S., & Donovan, T. (1994). Information and opinion change on ballot propositions. *Political Behavior*, *16*(4), 411-435.
- Brav, A., & Heaton, J. B. (2002). Competing theories of financial anomalies. *Rev. Financial Stud.*, *15*(2), 575-606.
- Brehmer, B. (1980). In one word: Not from experience. *Acta Psychologica*, *45*(1-3), 223-241.
- Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Rev.*, *78*, 1-3.
- Brown, S. D., & Steyvers, M. (2009). Detecting and predicting changes. *Cognitive Psychology*, *58*(1), 49-67.
- Bukszar, E., & Connolly, T. (1988). Hindsight bias and strategic choice: Some problems in learning from experience. *Acad. Management J.*, *31*(3), 628-641.
- Bush, R. R., & Mosteller, F. (1953). A stochastic model with applications to learning. *Ann. Math. Statist.*, *24*(4), 559-585.
- Camerer, C., & Ho, T.-H. (1999). Experience-weighted attraction learning in normal form games. *Econometrica*, *67*(4), 827-874.
- Chinnis, J. O., & Peterson, C. R. (1968). Inference about a nonstationary process. *J. Experiment. Psych.*, *77*(4), 620-625.
- Chinnis, J. O., & Peterson, C. R. (1970). Nonstationary processes and conservative inference. *J. Experiment. Psych.*, *84*(2), 248-251.
- Costa-Gomes, M., Crawford, V. P., & Broseta, B. (2001). Cognition and behavior in normal-form games: An experimental study. *Econometrica*, *69*(5), 1193-1235.
- Coursey, D. L., Hovis, J. L., & Schulze, W. D. (1987). The disparity between willingness to accept and willingness to pay measures of value. *Quart. J. Econom.*, *102*(3), 679-690.
- Cross, J. (1983). *A theory of adaptive learning economic behavior*. New York: Cambridge University Press.
- Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm*. Englewood Cliffs, NJ: Prentice-Hall.
- Dawes, R. M. (1979). Robust beauty of improper linear-models in decision-making. *Amer. Psychologist*, *34*(7), 571-582.
- Dawes, R. M., & Corrigan, B. (1974). Linear-models in decision-making. *Psych. Bull.*, *81*(2), 95-106.
- Deming, W. E. (1975). On probability as a basis for action. *Amer. Statistician*, *29*(4), 146-152.
- Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psych. Rev.*, *66*(3), 183-201.
- Ederer, F. P., & Manso, G. (2008). *Is pay-for-performance detrimental to innovation?* Available from [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1270384](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1270384)
- Einhorn, H., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psych. Rev.*, *85*(5), 395-416.

- Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *Amer. Econ. Review*, 88(4), 848-881.
- Estes, W. (1984). Global and local-control of choice behavior by cyclically varying outcome probabilities. *J. Experiment. Psych.: Learn. Memory, Cognition*, 10(2), 258-270.
- Fader, P. S., & Lattin, J. M. (1993). Accounting for heterogeneity and nonstationarity in a cross-sectional model of consumer purchase behavior. *Marketing Sci.*, 12(3), 304-317.
- Friedman, D. (1998). Monty Hall's three doors: Construction and deconstruction of a choice anomaly. *Amer. Econ. Rev.*, 88(4), 933-946.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psych. Rev.*, 103(4), 650-669.
- Gigerenzer, G., Todd, P. M., & The ABC Group. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psych.*, 24(3), 411-435.
- Grove, A. S. (1996). *Only the paranoid survive*. New York, NY: Doubleday.
- Heath, C., Larrick, R. P., & Klayman, J. (1998). Cognitive repairs: How organizational practices can compensate for individual shortcomings. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior* (Vol. 20, p. 1-37). Greenwich, CT: JAI Press.
- Hogarth, R. M. (2001). *Educating intuition*. Chicago, IL: University of Chicago Press.
- Hogarth, R. M., & Karelaia, N. (2005). Simple models for multiattribute choice with many alternatives: When it does and does not pay to face trade-offs with binary attributes. *Management Sci.*, 51(12), 1860-1872.
- Hogarth, R. M., & Karelaia, N. (2007). Heuristic and linear models of judgment: Matching rules and environment. *Psych. Rev.*, 114(3), 733-758.
- Hogarth, R. M., McKenzie, C. R. M., Gibbs, B. J., & Marquis, M. A. (1991). Learning from feedback: Exactness and incentives. *J. Experiment. Psych.: Learn. Memory, Cognition*, 17(4), 734-752.
- Kremer, M., Moritz, B., & Siemsen, E. (2011). Demand forecasting behavior: System neglect and change detection. *Management Science*, 57(Forthcoming).
- List, J. A. (2003). Does market experience eliminate market anomalies? *Quart. J. Econom.*, 118(1), 41-71.
- Lurie, N. H., & Swaminathan, J. M. (2009). Is timely information always better? the effect of feedback frequency on decision making. *Organ. Behav. Hum. Dec.*, 108(2), 312-329.
- Massey, C., & Wu, G. (2005a). Detecting regime shifts: The causes of under- and overreaction. *Management Sci.*, 51(6), 932-947.
- Massey, C., & Wu, G. (2005b). *Electronic companion paper: "Detecting regime shifts: The causes of under- and overreaction"*.
- McCullers, J. C. (1978). Issues in learning and motivation. In M. R. Lepper & D. Greene (Eds.), *The hidden costs of reward* (p. 5-18). Hillsdale, NJ: Erlbaum.
- McFadden, D. (1981). Econometric models of probabilistic choice. In C. F. Manski & D. McFadden (Eds.), *Structural analysis of discrete data with econometric applications* (p. 198-272). Cambridge, MA: MIT Press.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York: Cambridge University Press.
- Rapoport, A., Stein, W. E., & Burkheimer, G. J. (1979). *Response models for detection of change*. Dordrecht, Holland: D. Reidel.
- Robinson, G. H. (1964). Continuous estimation of a time-varying probability. *Ergonomics*, 7(1), 7-21.
- Rubin, M. A., & Girshick, H. (1952). A Bayes approach to a quality control model. *Ann. Math. Statist.*, 23(1), 114-125.
- Schweitzer, M. E., & Cachon, G. P. (2000). Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. *Management Sci.*, 46(3), 404-420.
- Shanteau, J. (1975). Averaging versus multiplying combination rules of inference judgment. *Acta Psychologica*, 39, 83-89.

- Shewhart, W. A. (1939). *Statistical method from the viewpoint of quality control*. Washington: Graduate School, Department of Agriculture.
- Sprecher, S. (1999). I love you more today than yesterday: Romantic partners' perceptions of changes in love and related affect over time. *J. Personality Soc. Psych.*, *76*(1), 46-53.
- Steineck, G., Helgesen, F., Adolfsson, J., Dickman, P. W., Johansson, J.-E., Norlen, B. J., et al. (2002). Quality of life after radical prostatectomy or watchful waiting. *New England J. Medicine*, *347*(11), 790-796.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.
- Theios, J., Brelsford, J. W., & Ryan, P. (1971). Detection of change in nonstationary binary sequences. *Perception Psychophysics*, *9*(6), 489-492.

## Appendix EXPERIMENTAL STIMULI

The stimuli for the experiment are shown below in Table A1, with “0” indicating a red ball for a particular observation, and “1” indicating a blue ball for that observation. Observations coming from the blue regime are indicated by bolded, underlined text. For example, bold, underlined text for observations 5 through 10 in sequence 1 indicate that the process changed in period 5, and that observations 1 through 4 were generated by the red regime and observations 5 through 10 were generated by the blue regime.

Trial	d	q	Period										Trial	d	q	Period									
			1	2	3	4	5	6	7	8	9	10				1	2	3	4	5	6	7	8	9	10
1	1.5	5%	1	0	0	0	<u>1</u>	<b>0</b>	<u>1</u>	<u>1</u>	<u>1</u>	<b>0</b>	21	1.5	20%	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>0</u>	<u>1</u>	
2	1.5	5%	0	0	1	1	1	1	1	0	0	1	22	1.5	20%	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	
3	1.5	5%	0	0	1	0	1	0	1	0	0	0	23	1.5	20%	0	1	0	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	
4	1.5	5%	0	1	1	1	0	0	0	1	1	0	24	1.5	20%	0	1	0	0	0	1	1	<u>0</u>	<u>0</u>	
5	1.5	5%	0	1	0	0	0	0	0	1	0	0	25	1.5	20%	<u>0</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>0</u>	<u>1</u>	<u>0</u>	
6	1.5	5%	0	1	1	0	0	1	0	0	0	1	26	1.5	20%	0	1	1	0	0	1	0	1	0	
7	1.5	5%	0	0	0	0	1	1	<u>0</u>	<u>1</u>	<u>0</u>	<u>0</u>	27	1.5	20%	0	0	0	0	1	0	1	0	0	
8	1.5	5%	1	1	0	1	1	1	0	0	1	0	28	1.5	20%	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>0</u>	
9	1.5	5%	0	1	1	0	0	1	1	0	1	0	29	1.5	20%	1	<u>0</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	
10	1.5	5%	0	0	0	1	1	0	1	<u>1</u>	<u>0</u>	<u>1</u>	30	1.5	20%	0	0	1	1	1	0	0	1	0	
11	1.5	5%	1	1	1	0	0	1	0	0	0	0	31	1.5	20%	0	1	1	0	1	0	1	0	0	
12	1.5	5%	1	0	0	0	0	0	0	1	1	1	32	1.5	20%	0	0	<u>0</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>0</u>	
13	1.5	5%	1	1	1	0	0	0	1	0	0	1	33	1.5	20%	0	<u>1</u>	<u>0</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>0</u>	<u>1</u>	<u>0</u>	
14	1.5	5%	0	0	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>0</u>	34	1.5	20%	0	0	0	0	1	0	0	1	0	
15	1.5	5%	1	1	1	0	1	1	0	1	0	1	35	1.5	20%	1	0	0	0	0	1	1	<u>1</u>	<u>0</u>	
16	1.5	5%	0	0	1	1	1	1	0	0	0	0	36	1.5	20%	<u>1</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>0</u>	<u>1</u>	<u>0</u>	
17	1.5	5%	0	0	0	1	0	0	0	0	1	1	37	1.5	20%	0	0	1	1	0	<u>0</u>	<u>1</u>	<u>1</u>	<u>0</u>	
18	1.5	5%	0	0	1	1	0	0	1	1	1	1	38	1.5	20%	0	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>0</u>	
19	1.5	5%	0	1	0	0	1	0	0	0	0	0	39	1.5	20%	0	0	<u>0</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>1</u>	
20	1.5	5%	0	1	0	1	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	40	1.5	20%	0	0	<u>0</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>0</u>	
41	3	5%	0	1	0	0	0	1	0	0	0	0	61	3	20%	0	1	0	0	0	1	0	<u>0</u>	<u>0</u>	
42	3	5%	0	0	1	0	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	62	3	20%	0	0	0	0	0	0	0	0	0	
43	3	5%	0	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	63	3	20%	0	1	0	0	1	0	0	0	0	
44	3	5%	0	0	0	1	1	0	1	0	0	0	64	3	20%	1	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	
45	3	5%	1	0	0	0	0	0	0	0	1	0	65	3	20%	0	0	0	1	0	0	0	1	<u>1</u>	
46	3	5%	0	0	0	0	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>0</u>	66	3	20%	0	1	0	0	0	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	
47	3	5%	1	1	0	0	0	0	0	0	0	0	67	3	20%	0	0	1	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	
48	3	5%	0	0	1	0	0	0	0	1	0	0	68	3	20%	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>0</u>	<u>1</u>	
49	3	5%	1	0	0	0	1	0	0	0	0	0	69	3	20%	0	0	1	0	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	
50	3	5%	<u>0</u>	<u>1</u>	<u>0</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	70	3	20%	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>		
51	3	5%	1	0	0	0	0	0	0	1	0	1	71	3	20%	0	0	0	1	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	
52	3	5%	0	0	0	0	0	0	1	0	0	0	72	3	20%	0	0	1	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	
53	3	5%	0	0	1	0	0	1	0	0	0	<u>1</u>	73	3	20%	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	
54	3	5%	0	0	0	0	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	74	3	20%	1	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>		
55	3	5%	0	0	0	0	0	1	1	0	1	0	75	3	20%	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	
56	3	5%	0	0	0	0	0	0	1	1	1	0	76	3	20%	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	
57	3	5%	0	0	0	0	0	0	0	1	0	1	77	3	20%	<u>1</u>	<u>1</u>	<u>0</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	
58	3	5%	0	0	1	0	0	0	0	0	0	0	78	3	20%	0	<u>1</u>	<u>1</u>	<u>0</u>	<u>0</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>1</u>	
59	3	5%	0	0	0	1	1	0	0	0	0	0	79	3	20%	1	0	0	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	
60	3	5%	1	0	0	1	0	0	<u>0</u>	<u>1</u>	<u>0</u>	<u>0</u>	80	3	20%	0	0	0	<u>1</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>0</u>	
81	9	5%	0	0	0	0	1	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	101	9	20%	0	1	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	
82	9	5%	0	0	0	0	0	0	0	0	0	0	102	9	20%	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	
83	9	5%	0	0	0	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	103	9	20%	1	0	0	0	0	1	0	0	<u>1</u>	
84	9	5%	0	0	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	104	9	20%	0	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	
85	9	5%	0	0	0	0	0	1	0	0	0	0	105	9	20%	0	0	0	1	0	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	
86	9	5%	0	0	0	0	0	0	0	0	0	0	106	9	20%	0	0	0	0	0	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	
87	9	5%	0	<u>0</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	107	9	20%	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	
88	9	5%	0	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	108	9	20%	0	0	0	1	0	0	<u>1</u>	<u>1</u>	<u>0</u>	
89	9	5%	0	0	0	0	0	1	1	0	0	0	109	9	20%	0	0	1	1	0	0	0	0	0	
90	9	5%	0	0	0	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	110	9	20%	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	
91	9	5%	0	0	0	0	0	0	0	0	0	0	111	9	20%	0	0	0	0	0	0	0	0	0	
92	9	5%	0	0	0	0	0	0	0	<u>1</u>	<u>1</u>	<u>1</u>	112	9	20%	0	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	
93	9	5%	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	113	9	20%	0	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>0</u>		
94	9	5%	0	0	0	0	0	0	0	<u>1</u>	<u>1</u>	<u>1</u>	114	9	20%	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	
95	9	5%	0	0	0	0	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	115	9	20%	0	0	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	
96	9	5%	0	0	0	0	0	0	0	0	1	0	116	9	20%	0	0	0	0	1	0	<u>1</u>	<u>1</u>	<u>1</u>	
97	9	5%	0	0	0	0	0	0	0	0	0	0	117	9	20%	0	0	0	0	1	0	0	<u>1</u>	<u>1</u>	
98	9	5%	1	0	0	0	0	1	0	0	0	0	118	9	20%	0	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	
99	9	5%	0	0	0	0	0	0	0	0	0	0	119	9	20%	0	0	0	1	0	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	
100	9	5%	0	0	0	0	0	0	1	0	0	0	120	9	20%	0	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	

Table A1: Stimuli from the learning experiment (0 indicates blue ball, 1 indicates red ball; bolded, underlined text indicates that signal was generated by the blue regime).