

Learning to Detect Change

Ye Li

Center for Decision Research, University of Chicago, Chicago, IL 60637, ye@chicagogsb.edu

Cade Massey

School of Management, Yale University, New Haven, IL 06511, cade.massey@yale.edu

George Wu

Center for Decision Research, University of Chicago, Chicago, IL 60637, wu@chicagogsb.edu

People, across a wide range of personal and professional domains, need to detect change accurately. Previous research has documented systematic shortcomings in doing so, in particular, a pattern of over- and underreaction to indications of change, resulting from a tendency to overweight signals of change at the expense of the environment that produces the signals. This investigation considers whether this pattern persists when participants are given the opportunity to learn. We find that the pattern of system neglect does persist, but that the impact of experience varies greatly across environments — participants show reliable improvement in some conditions and virtually none in others. We explain this differential learning by formally characterizing environments in terms of the extent to which they: (i) provide consistent feedback; and (ii) tolerate non-optimal behavior. Whereas we find that learning is related to consistent feedback, the stronger—and perhaps more surprising—finding is that more learning occurs in environments that are more tolerant of non-optimal behavior.

Key words: Change-point detection, regime shifts, learning, overreaction, underreaction, Bayesian updating.

1. 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. In the academic literature, the canonical example involves monitoring quality levels in a manufacturing process (Deming 1975, Rubin and Girshick 1952, Shewhart 1939). Recent research has broadened this paradigm beyond the field of operations. Researchers in finance have invoked regime-shifts to explain documented patterns of under- and overreaction in asset pricing (Barberis et al. 1998, Brav and Heaton 2002), while

economists have used change-point models to describe the challenges central bankers face in setting interest rates (Ball 1995, Blinder and Morgan 2005). However, relevant examples extend far beyond finance and economics: retailers must assess changes in consumer taste (Fader and Lattin 1993), corporate strategists must monitor technological trends in their marketplace (Grove 1996), politicians must track the sentiment of voters (Bowler and Donovan 1994), and individuals must pay attention to the health of their bodies (Steineck et al. 2002). And of course those with experience in romantic relationships understand the pitfalls of assessing a partner’s change in commitment or satisfaction (Sprecher 1999). Indeed, the need to detect change accurately is ubiquitous, and thus it is 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 the financial markets change from, say, a “bear” to a “bull” market. Economic indicators are at best imprecise signals, with the informativeness of these signals varying across indicators and over time. Underreacting to a signal of change may result in foregone opportunities to buy stocks at their lowest prices, while overreacting may result in owning still-declining shares.

A long literature in psychology investigates how successfully individuals navigate this task (e.g. Barry and Pitz 1979, Brown and Steyvers 2009, Chinnis and Peterson 1968, 1970, Estes 1984, Rapoport et al. 1979, Robinson 1964, Theios et al. 1971). A theme that emerges from this literature is that individuals respond to environmental conditions, but only partially. As Chinnis and Peterson (1968) stated: “[t]he 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 picked up on this theme, proposing that individuals react primarily to the signals they observe and secondarily to the environmental system that produced the signal. They tested this *system-neglect hypothesis* in experiments involving both probability estimation and choice. In these studies, 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 environment) were varied. In our investor example, these parameters correspond to the informativeness of market indicators and the historical rate of market vacillation, respectively. The experiments revealed a behavioral pattern consistent with the system-neglect hypothesis: underreaction was most common in unstable systems with precise signals, and overreaction was most prevalent in stable systems with noisy signals.

In the present paper, we consider how this pattern of system neglect is affected by experience. An oft-cited critique of behavioral decision research is that participants engage in relatively novel environments, without sufficient opportunity to learn (Coursey et al. 1987, List 2003). In Massey and Wu (2005a) (hereafter: MW), for example, participants changed environmental systems after each of the 18 trials. On the one hand, this arrangement enhances the salience of each environment's variables, increasing the likelihood that participants would give them sufficient attention. On the other hand, the design may have hindered participants' ability to adjust appropriately to these dimensions by minimizing their opportunity to learn about a particular environment. That design leaves open the issue of the robustness of the system neglect hypothesis. Does the systematic pattern of over- and underreaction observed in MW persist in the face of experience?

Whereas 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 configurations of diagnosticity or 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. Learning hinges on the whether feedback is informative, timely or uncertain, as well as the penalty for making inaccurate judgments. We take these two dimensions as useful for characterizing the conduciveness of environments toward learning.

Mapping feedback quality and the consequence of errors to the problem of change detection is not trivial. Whereas feedback quality is closely related to signal diagnosticity, the relationship of feedback quality to transition probability is not immediately clear. Even less clear is the relationship

between the consequence of errors and either signal diagnosticity or transition probability. Hence, we look beyond the system variables by mapping them into concrete operationalizations of feedback quality and the consequence of errors.

First we consider an 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 is optimal for other trials, or, in short, how clearly feedback points in the direction of optimal behavior. Based on extensive research showing that learning depends heavily on how relevant performance feedback is to future behavior (e.g. Einhorn and Hogarth 1978), we expect learning to be positively related to an environment’s consistency.

The second environmental factor we consider is *tolerance*, related to Hogarth’s consequence of errors. Tolerance reflects the extent to which a wide range of behavior produces payoffs similar to that generated by the optimal behavior, or, in short, how strong the incentives are to move in the direction of optimal behavior. 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, away from a deeper understanding of the task (Hogarth et al. 1991). While this can be beneficial in tasks requiring “simple, routine, unchanging responses and when circumstances favor the making of such responses quickly, frequently, and vigorously” (McCullers 1978), it is 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 performed worse than those under fixed-payment schemes. Given the complexity of our experimental task, we likewise expect more learning in 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 find evidence of positive relations between learning and both factors. This is in contrast to the mixed relation between learning and the system variables—we find some learning when signal

diagnosticity is high and little when it is low, but no consistent relation between learning and transition probability. Hence, by investigating consistency and tolerance, we relate the task of detecting change to the learning literature, shedding light on when and why learning occurs.

Our paper makes two main contributions. First, we extend the literature on change-point detection, establishing the robustness of the system-neglect hypothesis, along with the systematic pattern of under- and overreaction that is implied by our psychological hypothesis. Second, we develop and test explanations of learning in non-stationary settings. In so doing, we demonstrate the critical role of the environment in moderating the impact of experience on performance.

The outline of our paper is as follows. In

Section 2, we describe the statistical process used in our experimental studies. We then outline the system-neglect hypothesis in Section 3. Section 4 presents the design and results of our experimental study. We find strong evidence for system neglect in terms of the predicted pattern of over- and underreaction. In Section 5, we examine the relationship between decision environments and learning and find that environments that generate consistent feedback and that are tolerant to deviations from optimal behavior are more conducive to learning. We conclude by discussing implications of our findings and analyses for learning in other stochastic environments in Section 6.

2. Statistical Process

We begin by describing the statistical process used in our experimental study. The stochastic process consists of two regimes: the red the blue states. The stochastic process begins in the red state. The probability of a switch from the red state to the blue state in any period i is given by the transition probability, q . In addition, the blue state is an absorbing state (Feller 1968). Thus, the transition matrix is given by $P = \begin{bmatrix} 1 - q & q \\ 0 & 1 \end{bmatrix}$.

Participants observe either a red or blue signal signal in each period. A red signal is generated by the red state with probability $p_R > .5$ and by the blue state with probability $p_B < .5$. Thus, the process is a partially-observed Markov process (Monahan 1982). In our experiment, the probabilities are symmetric, $p_R = 1 - p_B$, and thus we take p_R/p_B to be a measure of the diagnosticity (d) of the signal, with larger values of d corresponding to more precise or diagnostic signals.

Let B_i (and R_i) indicate that the stochastic process is in the blue (red) state in period i . Furthermore, we denote the i th signal b_i , where $b_i = 1$ ($b_i = 0$) indicates that a blue (red) signal is observed in period i . Finally let $H_i = (b_1, \dots, b_i)$ denote the history of signals through period i . The Bayesian posterior odds of a change to the blue state after observing history H_i is then:

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

where $p_i^b = P(B_i|H_i)$ denotes the probability that the process has switched to the blue regime by period i .¹

3. The System-Neglect Hypothesis

The system-neglect hypothesis posits that participants will be more sensitive to signals than to system variables, thus resulting in a pattern of over- and underreaction like that observed in MW. This hypothesis builds on work by Griffin and Tversky (1992) in stationary environments (see also, Brenner et al. 2005). Griffin and Tversky argue 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 non-stationary 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 this hypothesis is that individuals are more likely to overreact in stable systems with noisy signals, and are more likely to underreact in unstable systems with precise signals. This prediction is captured schematically in Figure 1. Thus, the system-neglect hypothesis predicts two separate main effects, one of transition probability and one of diagnosticity. Note that the hypothesis makes a *relative* prediction and is silent about overall levels of reaction and thus is consistent with patterns of all underreaction or all overreaction.

¹ The derivation for Eqn. (1) is found in Massey and Wu (2005b). Note that the right hand side of Eqn. (1) factors out $(1 - (1 - q)^i)/(1 - q)^i$, the “base rate” odds of a change to the blue state in the absence of a signal.

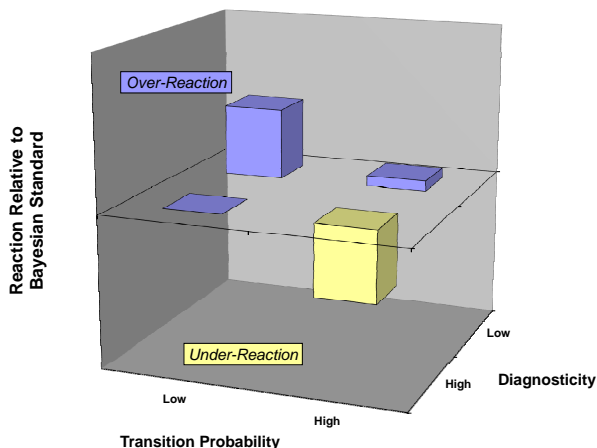


Figure 1 The system-neglect hypothesis predicts relatively more underreaction in precise, unstable systems and relatively more overreaction in noisy, stable systems.

To illustrate our hypothesis, consider four systems with two levels of diagnosticity, $d = 1.5$ and $d = 9$, crossed with two transition probabilities, $q = .05$ and $q = .20$. Suppose that the signals in the first two periods are both blue. The Bayesian posterior probabilities of a change to the blue state are $P(B_2|H_2) = .169$ when $d = 1.5$ and $q = .05$ (i.e., the noisy/stable system), whereas $P(B_2|H_2) = .920$ when $d = 9$ and $q = .20$ (i.e., the precise/unstable system). If individuals respond the same across the 4 conditions (for example, with a posterior probability of .60 in all conditions), 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 they be less responsive to diagnosticity and transition probability than to signals.

We also expect the hypothesized system-neglect pattern to attenuate over time, with judgments becoming more Bayesian with experience. However, we hypothesize that this learning will vary across conditions. Specifically, we expect that learning is most likely to occur when feedback is consistent and when the learning environment is tolerant.

4. Judgment Study

Our experimental setting allows us to compare individual judgments against the normative standard of Bayesian updating, as we provide participants in our study with all the information nec-

essary to calculate Bayesian responses and hence provide optimal judgments. Thus, our study is designed to test our system-neglect hypothesis by investigating how individuals revise probability judgments and whether these revisions move in the direction required by Bayesian updating.

4.1. Method

The study was conducted on computer using a specially designed Visual Basic program. We recruited 72 University of Chicago participants, with the task advertised as a “probability estimation task” in order to yield participants comfortable with probability. Participants were placed in one of 6 experimental conditions with 12 participants in each condition. The conditions were constructed by crossing 3 different diagnosticity levels ($d = 1.5, 3, \text{ and } 9$) with 2 different transition probability levels ($q = .05 \text{ and } .20$). The system variables remained constant for participants across the 20 trials. We will sometimes refer to these systems as noisy vs. somewhat precise vs. highly precise (for low, medium and high diagnosticities) and stable vs. unstable (for low and high transition probabilities). Our design differs from that employed in MW’s judgment study in several respects. MW used a total of 12 conditions, crossing the 3 diagnosticity conditions used in this study with 2 additional transition probabilities (.02 and .10). Participants in that study were also exposed to all 12 conditions. Whereas we have reduced that experimental design for simplicity, we have maintained the wide range of parameter values.

The computer program began by introducing the statistical process used in the experiment and illustrated the process using 4 demonstration trials and 2 practice trials. Following this introduction, each participant completed 20 trials consisting of 10 signals (periods). Participants were first shown the parameters, p_R , p_B , and q , governing their trials. These parameters were displayed continuously throughout each trial, and participants were told that the parameters would remain the same for all trials. They were then shown a sequence of red or blue signals drawn randomly based on that participant’s set of parameters. After seeing each signal, participants indicated the probability that the last signal was drawn from the blue state (i.e., the probability that the regime had already shifted). Participants were not permitted to change a probability once it was entered.

We randomly generated 20 unique sequences for each condition, using a statistical process with the true parameter values, yielding 120 sequences in total. Thus, each participant received the same 20 sequences as all other participants in that condition. The order of the sequences was randomized for each participant. The actual sequences are found in the electronic companion paper.

We compensated participants according to a quadratic scoring system that paid \$0.08 maximum (e.g., if a participant indicated with certainty that the process was in the blue state, and the process was in fact in the blue state) and \$-0.08 minimum (e.g., if a participant indicated with certainty that the process was in the blue state, and the process was in fact in the red state). Such a scheme is proper and thus truth-revealing for risk-neutral participants (Brier 1950). Whereas it was theoretically possible to lose money overall, doing so was exceedingly unlikely. To control for differences in difficulty across the different conditions, we provided participants in 3 of the 6 conditions with fixed payments in addition to the incentive payments. The fixed payments were designed so that a Bayesian decision maker would earn approximately the same amount in each condition. Fixed payments are found in row 3 of Table 1.

Participants were given feedback at the end of each trial as to if and when the process shifted from the red to the blue state. They were also informed how much money they made (or lost) on that particular trial as well as their cumulative earnings.

4.2. Results

In this section, we summarize the basic results of our empirical study. First we look at two measures of performance—earnings and the mean absolute difference between empirical and Bayesian judgments. Next, to test the system-neglect hypothesis, we consider measures of reaction—that is, *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 condition and then how they change with experience over the course of 20 trials.

4.2.1. Earnings Recall that participants were paid based on the difference between their subjective probability of a change to the blue state and the actual state (1 if blue, 0 if red). The

	Condition						Overall
	$d = 1.5$		$d = 3$		$d = 9$		
	$q = .05$	$q = .20$	$q = .05$	$q = .20$	$q = .05$	$q = .20$	
Empirical Earnings							
Variable Mean	10.25	6.77	11.57	10.68	13.48	13.10	10.97
Variable Std. Dev.	1.73	2.74	1.54	1.38	1.29	1.48	2.80
Fixed	2.00	5.00	0.00	1.00	0.00	0.00	
Total Mean	12.25	11.77	11.57	11.68	13.48	13.10	12.31
Bayesian Earnings							
Variable	12.44	9.55	13.12	13.14	14.58	14.36	12.86
Total	14.44	14.55	13.12	14.36	14.58	14.36	14.20
Difference between Bayesian and Empirical Mean Earnings							
All Quarters	2.19	2.78	1.56	2.46	1.10	1.26	1.89
Difference between Bayesian and Empirical Mean Earnings across Quarters (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
Empirical Mean Variable Earnings across Quarters (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

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), and empirical earnings by quarter. Data by quarter are scaled by a factor of 4.

mean of this absolute difference was .24 (median=.07, sd=.32), generating mean variable earnings of \$10.97 (range of \$-0.80 to \$14.48). The mean total earnings, including the fixed payment, was \$12.31 and ranged from \$4.20 to \$14.48. The mean earnings for each of the 6 conditions are presented in Table 1.

A Bayesian agent's mean earnings (i.e., a participant who provided probabilities according to

Bayes Rule) would have been \$12.86 for the variable portion and \$14.20 including the fixed payment. Thus, the mean deviation between Bayesian and empirical earnings was \$1.89 (range of \$0.08 to \$10.84 across participants). The earnings for a Bayesian agent for each of the 6 conditions are also found in Table 1. Participants tended to perform best, as measured by mean absolute deviation between Bayesian and empirical earnings, in the highly precise ($d = 9$) conditions, and worse in the noisy ($d = 1.5$) conditions. The deviations in the highly precise conditions are significantly lower than the deviations in the other conditions ($p < .02$). By the same measure, performance does not seem to depend on transition probability, with no difference between the stable ($q = .20$) and unstable ($q = .05$) conditions ($p > .20$).

To investigate the possibility of learning, we next consider how mean earnings change over the course of the 20 trials. We divided the 20 trials into 4 quarters of 5 trials per quarter. Table 1 lists the variable earnings by quarter while Figure 2 plots the difference between Bayesian and empirical earnings by condition and quarter (with lower values indicating better empirical performance). These values are re-scaled to be comparable to the overall variable earnings shown in Table 1 by multiplying the total earnings for each quarter by 4. Across all conditions there appears to be some learning, but it is most dramatic between Quarter 1 and 2, with performance improving in 5 of the 6 conditions (and for 42 of the 72 participants). Between Quarters 2 and 3, earnings increase in 4 of the 6 conditions (and for 33 of the 72 participants). Between Quarters 3 and 4 the mean performance actually falls, with earnings increasing in only 3 of the 6 conditions (and for 36 of the 72 participants).

Across conditions, the largest improvement occurs in the highly precise/stable condition ($d = 9$ and $q = .05$), with mean earnings per quarter increasing by \$0.44 from Quarter 1 to 4. In addition, mean earnings per quarter in Quarter 4 of that condition are only \$0.11 less than the Bayesian standard. The only other condition that produces steady improvement is the highly precise/unstable condition ($d = 9, q = .20$)

This analysis of earnings by quarter suggests there is learning in some conditions. We conduct

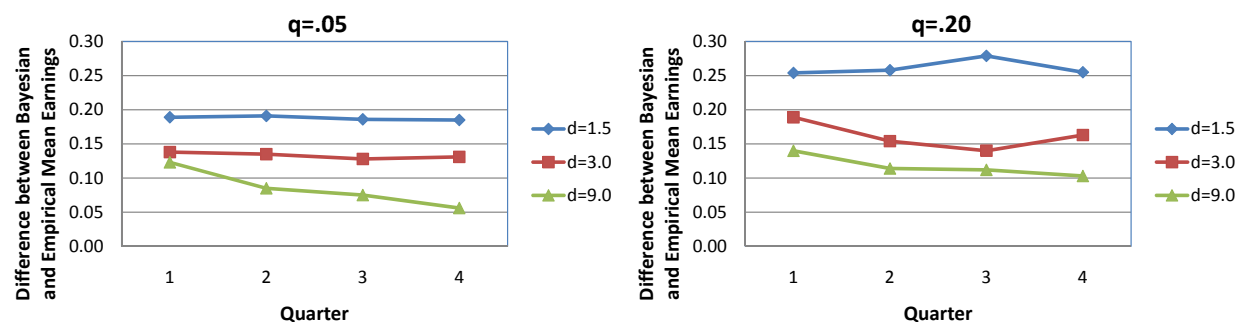


Figure 2 Learning by condition as measured by deviation between Bayesian and empirical earnings, with lower deviations indicating better performance. Note: Quarter earnings are rescaled by a factor of 4 to be comparable to earnings measures reported in Table 1.

Transition Probability	Diagnosticity		
	1.5	3	9
0.05	0.024 (0.045)	-0.004 (0.030)	0.098 (0.022)
0.20	0.036 (0.039)	0.014 (0.030)	0.036 (0.025)

Table 2 Change in variable earnings across trials for each condition, as measured by regression coefficients for trial, with mean earnings as the dependent variable. Standard errors are shown in parentheses.

Transition Probability	Diagnosticity		
	1.5	3	9
0.05	50%	50%	83%
0.20	58%	67%	50%

Table 3 Percentage of participants in each condition with improving earnings, as measured by a positive regression coefficient (regression takes mean earnings as the dependent variable and trial as predictor).

a stronger test by regressing the earnings by trial. To do so, we take an observation to be a trial-condition pair, i.e., each pair averages the 120 (12 participants per condition \times 10 periods per trial) earnings (again re-scaled to be comparable to overall earnings). Though Table 1 indicates that learning is not linear or even necessarily monotonic, for simplicity, we first assume linearity. The regression coefficients are found in Table 2. Coefficients in 5 of the 6 conditions are positive, though the improvement is significant only in the highly precise/stable condition. We also conducted the same analysis separately for each of the 72 participants. Results are shown by condition in Table 3. Coefficients are positive for 42 of the 72 (58%) participants ($p = .16$, sign test). Again, the only

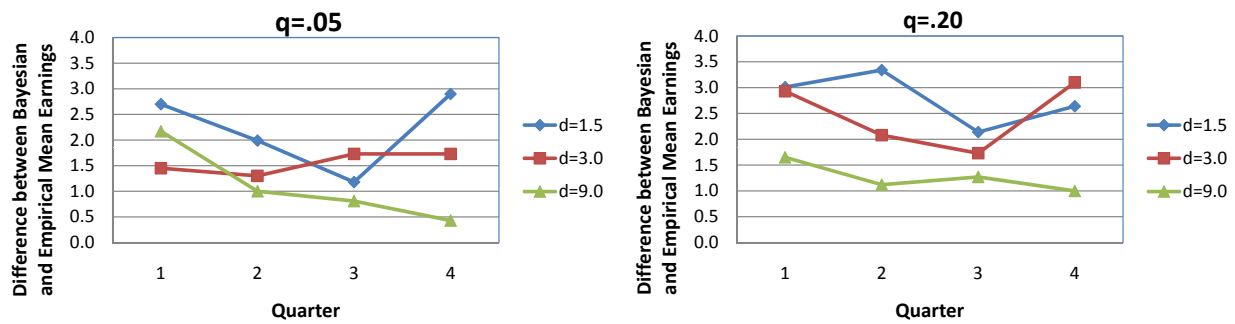


Figure 3 Learning by condition as measured by mean absolute deviation (MAD) between Bayesian and empirical probabilities by condition, with lower deviation indicating better performance.

significant improvement is shown in the $d = 9$, $q = .05$ condition, with 10 of the 12 participants having positive coefficients ($p = .02$, sign test).

4.2.2. Mean Absolute Deviations As a second measure of performance, we consider the mean absolute deviation (MAD) between empirical and Bayesian judgments. For series that are unrepresentative of the system that produces them, Bayesian judgments can produce lower earnings than non-Bayesian judgments. MADs are therefore attractive because they neither reward nor punish judgments based on luck and thus provide a less noisy performance measure.

The smallest MADs appear in the two highly precise conditions, with the largest MADs appearing in the two noisy conditions (see Table 4). Figure 3 plots this measure of performance across quarters. Overall, MADs decrease over the 4 quarters, though this pattern is not uniform across conditions. The largest and most consistent improvement appears again in the highly precise conditions, and in particular the highly precise/stable condition. There is also some learning in the moderately precise/unstable ($d = 3$, $q = .20$) condition (though mostly from Quarter 1 to 2), and virtually no learning in the other three conditions.

As in our analysis of earnings, we regressed the MADs across trials for each condition. Regression coefficients for this analysis are found in Table 5. The coefficients are negative (suggesting learning) for 5 of the 6 conditions, and significant in the two highly precise conditions and the moderately precise/unstable condition. The strongest learning again appears in the highly precise/stable condition.

	Condition						Overall
	$d = 1.5$		$d = 3$		$d = 9$		
	$q = .05$	$q = .20$	$q = .05$	$q = .20$	$q = .05$	$q = .20$	
Quarter 1	0.189	0.254	0.138	0.189	0.123	0.140	0.172
Quarter 2	0.191	0.258	0.135	0.154	0.085	0.114	0.156
Quarter 3	0.186	0.279	0.128	0.140	0.075	0.112	0.153
Quarter 4	0.185	0.255	0.131	0.163	0.056	0.103	0.149
All Quarters	0.188	0.261	0.133	0.161	0.085	0.117	0.158

Table 4 Mean absolute deviations (MADs) between empirical judgments and Bayesian judgments, by quarter and across conditions.

Transition Probability	Diagnosticity		
	1.5	3	9
0.05	-0.00066 (0.00091)	-0.00072 (0.00054)	-0.00400 (0.00077)
0.20	0.00040 (0.00112)	-0.00187 (0.00094)	-0.00235 (0.00060)

Table 5 Change in mean absolute deviations (MADs) between empirical and Bayesian judgments across trials for each condition, as measured by regression coefficients for trial, with MAD the dependent variable. Standard errors are shown in parentheses.

Transition Probability	Diagnosticity		
	1.5	3	9
0.05	42%	58%	92%
0.20	33%	75%	83%

Table 6 Percentage of participants in each condition with decreasing mean absolute deviations (MADs) between empirical and Bayesian judgments, as measured by negative regression coefficient (regression takes MAD as the dependent variable and trial as predictor).

We also ran regressions on each of the 72 participants separately (see Table 6). These regressions show a similar but stronger pattern to the analysis shown in Table 3. 46 of the 72 participants had negative coefficients ($p < .02$, sign test), with 11 of 12 participants improving in the highly precise/stable condition ($p < .01$, sign test). Otherwise, 10 of 12 participants improve in the highly precise/unstable condition ($p = .02$, sign test) and 9 of 12 participants in the moderately precise/unstable condition improve ($p = .08$, sign test). Overall, MADs are a more precise measure of performance than earnings, and thus we find somewhat stronger evidence of learning.

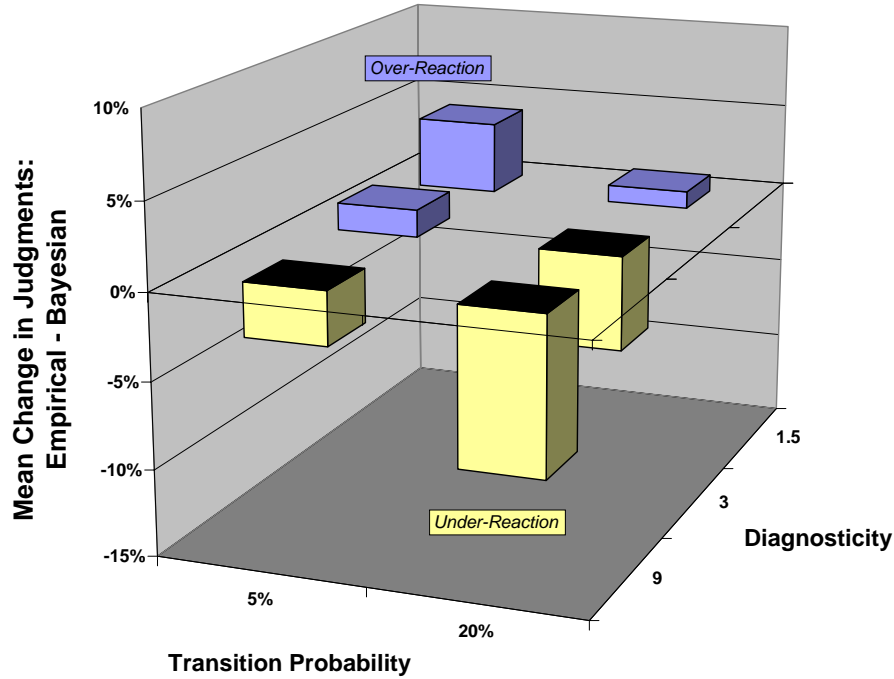


Figure 4 Over- and underreaction, by condition, as measured by the mean difference between the change in empirical probability judgments and the change in Bayesian probabilities, $\Delta p_i^e - \Delta p_i^b$.

4.2.3. Measures of reaction Whereas our analyses of earnings and MADs demonstrate learning differences across conditions, the system-neglect hypothesis specifies how empirical probability judgments respond to indications of change. Therefore, tests for system neglect require comparisons of *changes* in probabilities rather than absolute levels. To test for system neglect, we employ the same normative measure of change as used in MW. Let p_i^e be a participant's empirical probability of a change after observing i signals. The normative measure applies Bayes Rule to p_{i-1}^e :

$$\frac{\bar{p}_i^b}{1 - \bar{p}_i^b} = \frac{p_{i-1}^e}{1 - p_{i-1}^e} \left(\frac{1}{1 - q} \right) \left(\frac{p_R}{p_B} \right)^{2b_i - 1} + \left(\frac{q}{1 - q} \right) \left(\frac{p_R}{p_B} \right)^{2b_i - 1},$$

where \bar{p}_i^b is the Bayesian response to signal b_i , taking p_{i-1}^e as the “prior.” *Normative reaction* is then defined as $\Delta p_i^b = \bar{p}_i^b - p_{i-1}^e$. Note that this definition assumes that Bayes Rule should be applied regardless of the accuracy of the prior. We take this approach in order to focus on belief *revision*, at each stage granting participants their priors and evaluating only how beliefs are updated.

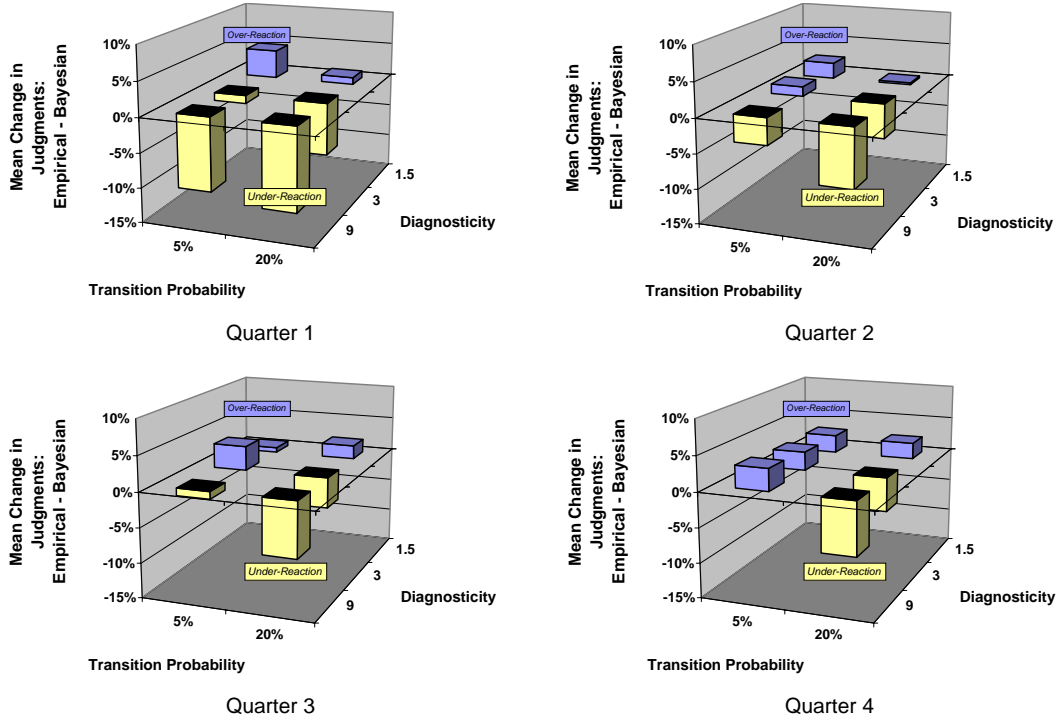


Figure 5 Over- and underreaction, by condition and quarter, as measured in Figure 4.

We compare the normative reaction to the *empirical reaction*, $\Delta p_i^e = p_i^e - p_{i-1}^e$, with underreaction indicated by $\Delta p_i^e < \Delta p_i^b$ (or $p_i^e < \bar{p}_i^b$) and overreaction indicated by $\Delta p_i^e > \Delta p_i^b$ (or $p_i^e > \bar{p}_i^b$). The system-neglect hypothesis predicts a greater tendency to underreact in more precise, less stable conditions, and to overreact in noisier, stabler conditions. Figure 4 depicts the mean error in reaction, $\Delta p_i^e - \Delta p_i^b$, in our 6 experimental conditions.² As predicted by the system-neglect hypothesis, the greatest underreaction occurs in the southeast-most cell ($d = 9, q = .20$), while the greatest overreaction occurs in the northwest-most cell ($d = 1.5, q = .05$). For all 12 of the pairwise comparisons, the pattern of underreaction is monotonic as transition probability and diagnosticity increase.

Figure 5 plots the mean error in reaction for each quarter. Note that the degree of system neglect is most pronounced in Quarter 1, but remains significant in the remaining three quarters. Consistent with the analyses in the previous sections, it appears that the learning that does take

² Since the predicted gradient applies only to indications of change, Figure 4 only includes reactions to blue signals. Red signals can be interpreted as signals of “non-change” and are predicted to exhibit a different gradient. See Massey and Wu (2005a), pp. 934-935.

place occurs mostly in the highly precise conditions, and from Quarter 1 to Quarter 2.

4.2.4. Estimation Although the pattern of reactions in Figure 4 is consistent with the system-neglect hypothesis, it is nevertheless possible that the hypothesized pattern could be an artifact of the specific random sequences. We formally rule out this alternative explanation by estimating a *quasi-Bayesian* model to test for system neglect (see MW for a more complete discussion of this model). The *quasi-Bayesian* model is a generalization of Eqn. (1) that explicitly allows for non-optimal sensitivity to transition probability and diagnosticity. We do so by adding two parameters, α and β , to the Bayesian expression in Eqn. (1):

$$\frac{p_i^e}{1-p_i^e} = \frac{\Pr(B_i|H_i)}{\Pr(R_i|H_i)} = \left(\frac{1-(1-\alpha q)^i}{(1-\alpha q)^i} \right) \sum_{j=1}^i \frac{q(1-q)^{j-1}}{1-(1-q)^i} d^\beta \left[i+1-j - \binom{i}{k=j} b_k \right]. \quad (2)$$

In Eqn. (2), α moderates sensitivity to transition probability, and β moderates sensitivity to diagnosticity. Note that $\alpha < 1$ and $\beta < 1$ reflect insensitivities to transition probability and signals, respectively. In addition, $p_i^e \rightarrow 0$ as $\alpha \rightarrow 0$ and $p_i^e/(1-p_i^e) \rightarrow (1-(1-\alpha q)^i)/(1-\alpha q)^i$ (the odds of change in the absence of any signals) as $\beta \rightarrow 0$. We return the Bayesian expression (Eqn. (1)) when $\alpha = \beta = 1$.

The system-neglect hypothesis predicts that parameter values differ systematically across environments; however, Eqn. (2) explicitly specifies that parameter values are the same in all environments. Constant parameter values are actually quite sensitive to the environment, serving only to monotonically modify q and d . Thus, whereas Eqn. (2) can measure the overall level of conservatism (cf. Edwards 1968), it must be further generalized to test for system neglect. Formally, let $\alpha = \alpha_1 Q_1 + \alpha_2 Q_2$ and $\beta = \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3$, where Q_m and D_n are indicator variables corresponding to transition probability q_m and diagnosticity d_n , respectively. System neglect requires the following monotonic ordering of these parameters: $\alpha_1 > \alpha_2$ for $q_1 = .05 < q_2 = .20$ and $\beta_1 > \beta_2 > \beta_3$ for $d_1 = 1.5 < d_2 = 3 < d_3 = 9$. To see why this ordering is required, consider the extreme case in which individuals neglect the system completely and only react to signals. In this case of *complete* system neglect, $q_1 \times \alpha_1 = q_2 \times \alpha_2$ and $d_1^{\beta_1} = d_2^{\beta_2} = d_3^{\beta_3}$. Even though we do not predict such extreme

system neglect, we nevertheless predict the monotonic ordering specified above. Note that the system-neglect hypothesis does *not* specify the overall level of conservatism, and this formulation thus allows for both conservatism ($\alpha_m < 1$ or $\beta_n < 1$) and “radicalism” ($\alpha_m > 1$ or $\beta_n > 1$).

We first run the model by pooling all the data across individuals and conditions and estimating p_i^e using Eqn. (2), in a single omnibus nonlinear regression.³ In so doing, we model the behavior of a representative agent, an approach often used in econometrics (McFadden 1981). Figure 6 shows the estimates of α_m and β_m , where lower levels imply greater conservatism. The estimated parameters are ordered as predicted by system neglect: In all cases, $\alpha_m > \alpha_n$, and $\beta_m > \beta_n$, where $m > n$, with all pairwise differences significant at $p < .001$. In addition, there is a mixed pattern of conservatism and radicalism in both parameters, consistent with the pattern depicted in Figure 4. It is also useful to compare the gradients of the transition probability and diagnosticity estimates with comparable curves for complete system neglect. These curves are normalized to coincide with α_1 and β_1 , respectively. Note that whereas the slopes of the transition probability and diagnosticity parameters are much steeper than Bayesian, they are less shallow than the slope of the complete-neglect curves. In other words, though participants are not sufficiently sensitive to system variables, they are not completely *insensitive* to them, either. We also separately estimated Eqn. (2) for each individual. A plot of the means of these estimates across conditions is qualitatively similar to Figure 6.

We further investigate whether learning is responsible for the observed differences in system neglect by looking at estimates by quarter, again dividing the data into 4 quarters of 5 trials each. We estimate parameters for each quarter, using the same representative agent approach with additional dummies for quarter. Figure 7 shows the estimates for all 4 quarters. There are several inferences to draw from this figure. First, learning occurs for stable ($q = .05$) conditions but not unstable ones ($q = .20$). Second, the parameter estimates for both moderately ($d = 3$) and highly

³ We cannot perform estimates on either means or medians because we are concerned with learning across trials, and participants see sequences in random order. Summary statistics within trial would necessarily be *across* sequence. The additional noise generated by different orderings of sequences makes any summary statistics inappropriate for estimation.

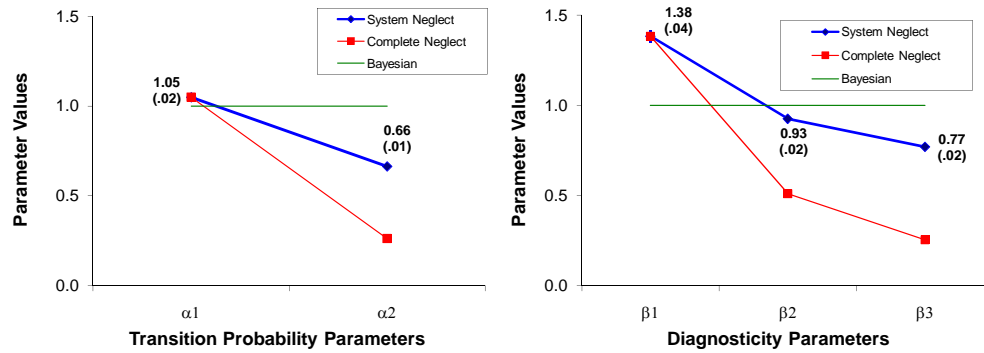


Figure 6 System neglect parameter estimates by condition

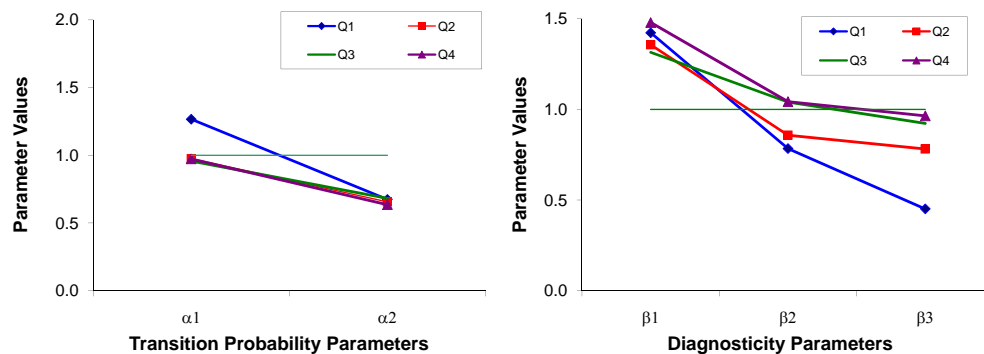


Figure 7 System neglect parameter estimates by quarter and condition

precise ($d = 9$) conditions appear to converge toward the Bayesian standard of 1. This convergence is particularly dramatic for the highly precise conditions. In contrast, there is virtually no learning in the noisy conditions ($d = 1.5$). Third, most of the learning occurs early, between Quarters 1 and 2. These results are consistent with the analyses conducted on earnings and MADs.

The analysis presented in this section revealed that: (i) learning tends to occur early, if at all; and (ii) learning is most prevalent in the highly precise/stable condition ($d = 9, q = .05$), followed by the highly precise/unstable condition ($d = 9, q = .20$). There is modest learning in the moderately precise/unstable condition ($d = 3, q = .20$) and hardly any learning in any of the other three conditions. The system-neglect hypothesis is silent on why learning varies in this way across experimental conditions. In order to better understand this variation, we turn to a more direct analysis of the learning environment in the next section.

5. Learning Environment

We have analyzed learning thus far in terms of our system variables, diagnosticity and transition probability. However, it is more informative to more directly examine the learning environments faced by our participants. In this section we investigate two environmental factors that critically affect how we learn: the quality of feedback (consistency) and the consequence of errors (tolerance). We first examine how these factors vary as a function of the system variables, and then explore their impact on learning. Our analysis thus 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 research on learning in stochastic environments that directly examines a reinforcement learning process (Bush and Mosteller 1953, Camerer and Ho 1999, Cross 1983, Erev and Roth 1998). We adopt our modeling approach for two reasons. First, just as we focus on environmental factors to understand when system neglect leads to under- and overreaction, we focus on environmental factors to understand what leads to changes in system neglect over time. Second, the environments we study are dynamic. As such, responses are meaningful only in context. For example, a response of 50% on the first observation of a trial is very different from a response of 50% after drawing three consecutive blue balls. 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 α and β toward optimal parameter values. While such an approach seems reasonable, it would be quite a departure from previous research.

5.1. Environmental factors and experimental design

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

To operationalize feedback consistency, we denote $Q_t(\alpha_t, \beta_t)$ the earnings in trial t from responding in a quasi-Bayesian fashion with α_t and β_t (see Eqn. (2)). Also let α_t^* and β_t^* be the parameters that maximize earnings in trial t . A challenge in using this definition is that many trials do not have unique optima. For example, β_t^* is not uniquely determined in trials with no red signals and neither β_t^* nor α_t^* are determined in trials in which there are no blue signals. To address this problem, we define α_t^* and β_t^* as the midpoint of the optimal plateau for these trials. Then, feedback is more *consistent* for large values of

$$\frac{1}{19 \times 20} \sum_t \sum_{t' \neq t} \frac{Q_{t'}(\alpha_t^*, \beta_t^*)}{Q_t(\alpha_t^*, \beta_t^*)}.$$

This measure reflects the mean degree to which the optimal quasi-Bayesian parameters in one trial are reinforced in later trials, relative to the maximum earnings achievable on those trials. (The mean matches each of the 20 trials per condition with the other 19 trials in that condition.)

Second, we consider the consequence of inaccurate judgments. An environment is more *tolerant* if errors are not punished very harshly relative to correct judgments. Tolerance is operationalized as the degree to which incorrect α and β parameters generate earnings within a certain threshold of the maximum earnings. In other words, more tolerant environments have flatter maxima. Let $R(\alpha, \beta)$ denote the earnings from using parameters α and β for the 20 series in a particular condition, with α^{**} and β^{**} denoting the parameters that maximize earnings for that condition, summed across the 20 trials. For each of our 6 conditions, we calculate earnings resulting from all combinations of $0 \leq \alpha \leq 5$ and $0 \leq \beta \leq 5$. We choose to cap the upper range of each variable at 5 due to diminishing marginal effects of higher values on judgment beyond this point. We then determine what percentage of combinations of α and β produce earnings within some threshold τ of the maximum earnings, i.e., $R(\alpha^{**}, \beta^{**}) - R(\alpha, \beta) \leq \tau$.

Before turning to an analysis of learning, we first consider how consistency and tolerance relate to our system variables. Table 7 shows the consistency measure for each condition. Feedback is most consistent in the highly precise ($d = 9$) conditions, followed by the noisy/stable condition ($d = 1.5$, $q = .05$). It is least consistent in the noisy/unstable condition ($d = 1.5$, $q = .20$) A linear regression

of consistency on diagnosticity, transition probability, and their interaction reveals that stable systems are more consistent than unstable ones ($p < .05$). This effect is attenuated for moderately precise ($d = 3$) conditions and slightly reversed for highly precise ones ($p < .05$).

Table 7 shows tolerance measures for each condition for τ ranging from \$0.50 to \$2.00. This analysis shows that the earning space is remarkably flat for the highly precise conditions, and relatively flat for the moderately precise/unstable condition ($d = 3, q = .20$). It is quite steep for the other two stable conditions ($d = 1.5, q = .05$ and $d = 3, q = .05$). For example, 50% of the combinations of α and β for the highly precise/stable ($d = 9, q = .05$) condition perform within \$1.50 (roughly 90%) of the maximum earnings. In contrast, the same measure is less than 5% for the other two stable conditions. Figure 8 shows this measure of tolerance graphically. We see that both highly precise conditions have flat maxima with a large region of α and β producing near optimal earnings. In contrast, both noisy conditions have very small regions that produce near-optimal earnings. A regression of tolerance on diagnosticity, transition probability, and their interaction reveals only a positive main effect of diagnosticity ($p < .05$). All coefficients have the same sign across all levels of tolerance. The main takeaway is that environments with highly precise signals are highly tolerant to error.

In general we see that these environmental factors are influenced by the system variables, but in complicated ways. It is precisely because of this complexity that consistency and tolerance may relate more strongly to learning than do the system variables.

5.2. Environmental factors and learning

Next, we relate these two environmental factors to the learning we observe across conditions. To do so, we use the measure of learning reported in Table 5, the change in mean absolute deviation (MAD) over time (MAD learning, for short). Recall that this measure was determined by regressing the mean absolute deviations between empirical judgments and Bayesian posteriors on trial. To facilitate comparison, we repeat these coefficients in the bottom row of Table 7.

In our analyses we focus on a moderate level of tolerance, the percentage of parameter values within \$1.50 of the optimum (row 6 in Table 7). This threshold corresponds to roughly 90% of the

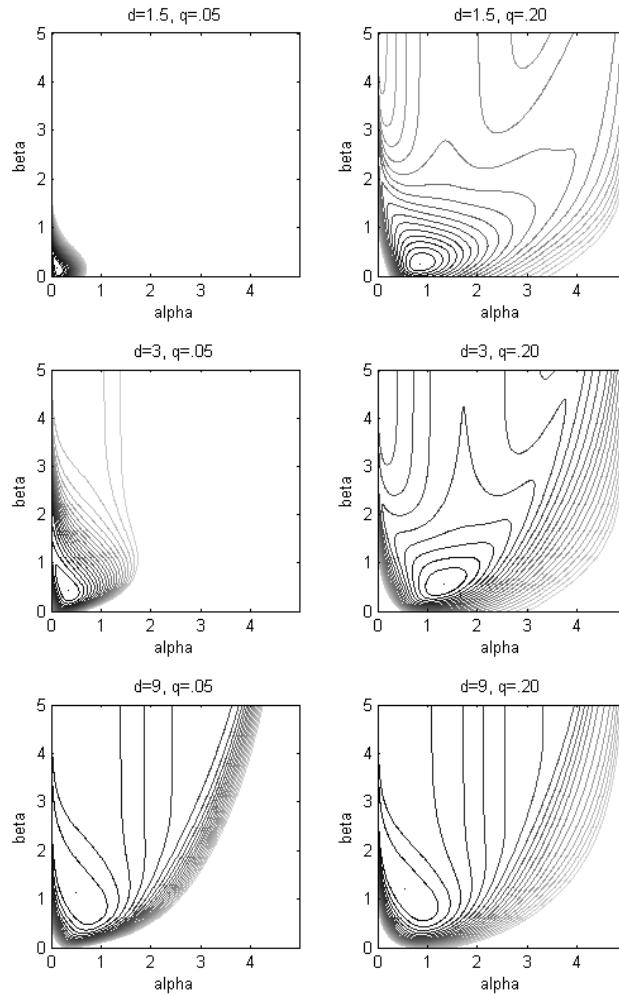


Figure 8 Earnings contours assuming quasi-Bayesian model. For each condition, the plots show the total earnings across the 20 trials for all combinations of $0 \leq \alpha \leq 5$ and $0 \leq \beta \leq 5$. The contours are spaced at 25¢ intervals from the maximum.

maximum earnings. We choose this level for two empirical reasons—“narrower” levels of tolerance are more highly correlated with consistency, and the relation between environmental factors and system variables is strongest at this level (as seen in the R^2 of the following regressions compared to those with “narrower” and “wider” thresholds). There is no *a priori* correct threshold, so we check the robustness of our results using alternative specifications, reporting inconsistencies.

Figure 9 plots the measure of learning as a function of consistency and tolerance. The figure reveals a strong relationship between learning and tolerance: in 13 of the 15 pairwise comparisons, learning is higher when tolerance is higher. There is also a relationship between learning and

	Condition					
	$d = 1.5$		$d = 3$		$d = 9$	
	$q = .05$	$q = .20$	$q = .05$	$q = .20$	$q = .05$	$q = .20$
Consistency (Mean performance using optima of other series)	61.5%	18.0%	53.1%	34.2%	66.8%	73.8%
Tolerance (Percent of parameter space within τ 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%
Learning (Per trial improvement in MAD)						
	-.00066	.00040	-.00072	-.00187	-.00400	-.00235

Table 7 Analysis of learning environments. The top row measures consistency, indicating the mean performance of each trial's optimal parameters on all other trials. Higher values indicate more consistent feedback. The middle rows measure tolerance, indicating what percentage of the $0 \leq \alpha \leq 5$ and $0 \leq \beta \leq 5$ space produces total earnings within some threshold for a condition of the condition-maximizing parameters, α^{**} and β^{**} . Higher values indicate more tolerant environments. The bottom row shows the amount of learning observed in each condition, as measured by the change in mean absolute deviation (MAD) across trials. More negative values reflect more learning.

consistency, though a somewhat weaker one: in 11 of the 15 pairwise comparisons across conditions, learning is higher when consistency is higher.

We look more formally at these relationships by regressing MAD learning on our measures of consistency and tolerance. We reverse score MAD learning so that positive coefficients correspond to more learning. Since our sample size is exceedingly small ($n = 6$) and our dependent variable is estimated with error, results should be interpreted with caution.

Consistent with the informal analysis, tolerance has a significant positive effect on learning ($t = 4.75$, $p = .01$) while consistency has a marginally positive effect ($t = 2.59$, $p = .08$). We replicate this regression using other thresholds of tolerance at \$0.10 intervals up to \$3.00. The effects of tolerance and consistency are always in the same direction, though the results for consistency

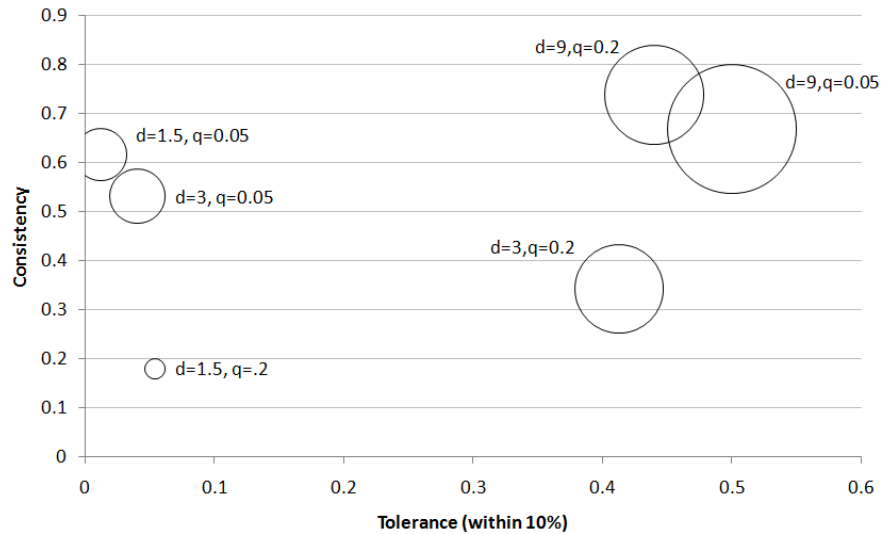


Figure 9 Learning across different conditions as a function of consistency and tolerance. The diameter of the circle is proportional to the amount of learning as indicated by the reverse scored regression coefficient of MAD on trial. Consistency is measured by the mean performance using each trial's optima 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.

become insignificant for thresholds less than \$1.40.⁴

A positive relationship between consistency and learning is not surprising. There may be a number of reasons the present analysis reveals only a marginally significant relationship. In addition to the limited sample size and measurement error in both the independent and dependent variables, our environments may not span a broad enough range of consistency. Highly consistent feedback—beyond that which we evaluate—likely enables 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

⁴ We also ran a second regression which included the interaction between consistency and tolerance. Whereas no coefficients change signs, there is a marked change in significance level. In a reversal, the “narrower” thresholds show significant results on all three terms. For example, for the \$0.50 threshold, consistency ($t = 6.24$, $p < .05$) and tolerance ($t = 5.60$, $p < .05$) are both positively related to learning, while the coefficient on the interaction term is negative ($t = 4.67$, $p < .05$). For “wider” thresholds, tolerance and the interaction term lose significance. Again, we must interpret these results with extreme caution due to the low ratio of data points to independent variables.

it. In the language of the contour maps presented in Figure 8, feedback points the direction to the (local) maxima, while incentives indicate the potential change in elevation achieved by moving in that direction. The negative relation between incentives (i.e., less tolerance) and learning indicate that, holding feedback constant, individuals learn better with flat maxima. That is, they improve performance more when they have the least 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 argue 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 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. While we are cautious in generalizing too much from this observation, it is provocative in 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 environment. Drawing on research in psychology, we find that environments vary significantly in their conduciveness to learning. The pattern is stark, despite a small sample and sparse stimuli: learning depends on the tolerance of the environment and on the consistency of feedback.

6. General Discussion

This research was motivated primarily by the desire to pit system neglect against learning. Given sufficient experience in an environmental “system,” would participants grow more sensitive to its key features, learning to avoid the over- and underreaction driven by system neglect? The answer is, in general, no. Across a wide range of environments participants display 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. The participants gained experience sufficient

for some learning in some conditions. However, where we observe learning, it was not enough to overcome system neglect, and in many conditions we observe no learning at all.

This observation underscores both the robustness of system neglect and the challenge of learning. Experience alone is not sufficient to overcome judgmental biases. Rather, the impact of experience is moderated by one's environment. In our experimental study, we see suggestions of learning across most conditions, but reliable evidence in only 2 or 3 of the 6 conditions. Learning to avoid over- and underreaction depends somewhat on the consistency of feedback and strongly on the tolerance of the environment to errors. Whereas we see modest improvements in the early trials of most of the conditions, this improvement persists only in highly tolerant environments or in moderately tolerant environments with highly consistent feedback. That is, consistent with an argument developed by Hogarth et al. (1991), we find that stronger incentives for improved performance actually impair learning.

The implications of this research extend beyond change-point detection. Below we discuss the contributions to research on learning in stochastic environments and organizational decision-making.

6.1. 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 et al. 1951). In this problem, the retailer must balance the cost of overstocking (the unit cost of the item minus some salvage costs) and the cost of understocking (the lost profit and possibly some 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. Nevertheless, many empirical studies have shown relatively poor performance and modest learning at best (Schweitzer and Cachon 2000, Bolton and Katok 2008).

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

distribution leads to lower consistency of feedback. And 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 and 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 study 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.

6.2. 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 and March 1963). Another approach is the use of policies, restricting the freedom decision-makers have (Heath et al. 1998). One of the goals of such programs is to keep decision makers from "chasing noise," 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

and/or short-term incentives, decrease tolerance by creating an environment that is very responsive to success and failure (Baron and Hershey 1988, Bukszar and Connolly 1988). These incentives take many forms, from pay-for-performance and promotion to the way failure is handled informally. 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 (Ederer and 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 incentive intensity and learning.

6.3. Conclusion

Overall, our main contributions are to establish the robustness of system neglect in change-point detection and to demonstrate the relationship between environmental conditions and learning. In the end, we are somewhat sober about the ability of individuals to avoid systematic under- and overreaction 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

- Arrow, K. J., T. Harris, J. Marschak. 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., A. Shleifer, R. Vishny. 1998. A model of investor sentiment. *J. Financial Econom.* **49**(3) 307–343.
- Baron, J, J. Hershey. 1988. Outcome bias in decision evaluation. *J. Personality Soc. Psych.* **54**(4) 569–579.
- Barry, D. M., G. F. Pitz. 1979. Detection of change in nonstationary, random sequences. *Organizational Behavior and Human Performance* **24**(1) 111–125.
- Blinder, A. S., J. Morgan. 2005. Are two heads better than one? Monetary policy by committee. *J. Money, Credit, and Banking* **37**(5) 789–811.

- Bolton, G. E., E. Katok. 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., T. Donovan. 1994. Information and opinion change on ballot propositions. *Political Behavior* **16**(4) 411–435.
- Brav, A., J. B. Heaton. 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.
- Brenner, L., D. Griffin, D.J. Koehler. 2005. Modeling patterns of probability calibration with random support theory: Diagnosing case-based judgment. *Organ. Behav. Human Decision Processes* **97**(1) 64–81.
- Brier, G. W. 1950. Verification of forecasts expressed in terms of probability. *Monthly Weather Rev.* **78** 1–3.
- Brown, S. D., M. Steyvers. 2009. Detecting and predicting changes. *Cognitive Psychology* **58**(1) 49–67.
- Bukszar, E., T. Connolly. 1988. Hindsight bias and strategic choice: Some problems in learning from experience. *Acad. Management J.* **31**(3) 628–641.
- Bush, R. R., F. Mosteller. 1953. A stochastic model with applications to learning. *Ann. Math. Statist.* **24**(4) 559–585.
- Camerer, C., T.-H. Ho. 1999. Experience-weighted attraction learning in normal form games. *Econometrica* **67**(4) 827–874.
- Chinnis, J. O., C. R. Peterson. 1968. Inference about a nonstationary process. *J. Experiment. Psych.* **77**(4) 620–625.
- Chinnis, J. O., C. R. Peterson. 1970. Nonstationary processes and conservative inference. *J. Experiment. Psych.* **84**(2) 248–251.
- Coursey, D. L., J. L. Hovis, W. D. Schulze. 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*. Cambridge University Press, New York.
- Cyert, R. M., J. G. March. 1963. *A behavioral theory of the firm*. Prentice-Hall, Englewood Cliffs, NJ.
- 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., G. Manso. 2008. Is pay-for-performance detrimental to innovation? URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1270384.
- Edwards, W. 1968. Conservatism in human information processing. B. Kleinmuntz, ed., *Formal representation of human judgment*. Wiley, New York, 17–52.
- Einhorn, H.J., R. M. Hogarth. 1978. Confidence in judgment: Persistence of the illusion of validity. *Psych. Rev.* **85**(5) 395–416.
- Erev, I., A. E. Roth. 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.K. 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., J. M. Lattin. 1993. Accounting for heterogeneity and nonstationarity in a cross-sectional model of consumer purchase behavior. *Marketing Sci.* **12**(3) 304–317.
- Feller, W. 1968. *An introduction to probability theory and its applications*. 3rd ed. Wiley, New York.
- Friedman, D. 1998. Monty Hall's three doors: Construction and deconstruction of a choice anomaly. *Amer. Econ. Rev.* **88**(4) 933–946.
- Griffin, D., A. Tversky. 1992. The weighing of evidence and the determinants of confidence. *Cognitive Psych.* **24**(3) 411–435.
- Grove, A. S. 1996. *Only the paranoid survive*. Doubleday, New York, NY.
- Heath, C., R. P. Larrick, J. Klayman. 1998. Cognitive repairs: How organizational practices can compensate for individual shortcomings. B. M. Staw, L. L. Cummings, eds., *Research in organizational behavior*, vol. 20. JAI Press, Greenwich, CT, 1–37.
- Hogarth, R. M. 2001. *Educating intuition*. University of Chicago Press, Chicago, IL.
- Hogarth, R. M., C. R. M. McKenzie, B. J. Gibbs, M. A. Marquis. 1991. Learning from feedback: Exactingness and incentives. *J. Experiment. Psych.: Learn. Memory, Cognition* **17**(4) 734–752.
- List, J. A. 2003. Does market experience eliminate market anomalies? *Quart. J. Econom.* **118**(1) 41–71.
- Massey, C., G. Wu. 2005a. Detecting regime shifts: The causes of under- and overreaction. *Management Sci.* **51**(6) 932–947.

- Massey, C., G. Wu. 2005b. Electronic companion paper: “Detecting regime shifts: The causes of under- and overreaction”.
- McCullers, J. C. 1978. Issues in learning and motivation. M. R. Lepper, D. Greene, eds., *The Hidden Costs of Reward*. Erlbaum, Hillsdale, NJ, 5–18.
- McFadden, D. 1981. Econometric models of probabilistic choice. C. F. Manski, D. McFadden, eds., *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press, Cambridge, MA, 198–272.
- Monahan, G. E. 1982. A survey of partially observable markov decision processes: Theory, models, and algorithms. *Management Sci.* **28**(11) 1–16.
- Rapoport, A., W. E. Stein, G. J. Burkheimer. 1979. *Response models for detection of change*. D. Reidel, Dordrecht, Holland.
- Robinson, G. H. 1964. Continuous estimation of a time-varying probability. *Ergonomics* **7**(1) 7–21.
- Rubin, M. A., H. Girshick. 1952. A bayes approach to a quality control model. *Ann. Math. Statist.* **23**(1) 114–125.
- Schweitzer, M. E., G. P. Cachon. 2000. Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. *Management Sci.* **46**(3) 404–420.
- Shewhart, W. A. 1939. *Statistical Method from the Viewpoint of Quality Control*. Graduate School, Department of Agriculture, Washington.
- 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., F. Helgesen, J. Adolfsson, P. W. Dickman, J.-E. Johansson, B. J. Norlen, L. Holmberg. 2002. Quality of life after radical prostatectomy or watchful waiting. *New England J. Medicine* **347**(11) 790–796.
- Sutton, R. S., A. G. Barto. 1998. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA.
- Theios, J., J. W. Brelsford, P. Ryan. 1971. Detection of change in nonstationary binary sequences. *Perception Psychophysics* **9**(6) 489–492.