Learning to Detect Change

Ye Li[1]
Cade Massey[2]
George Wu[3]

May 17, 2013

ABSTRACT

People need to accurately detect change across a wide range of personal and professional domains. Previous research has documented a systematic pattern of over- and under-reaction to signals of change because of system neglect, or the tendency to overweight signals and underweight the system producing the signals. We investigate whether experience improves ability to detect change. Participants in our study made judgments in 20 trials of 10 periods per trial, all in a single system. We found that although the system-neglect pattern is not completely attenuated by experience, average performance did improve with experience. However, the degree of learning varied substantially across the 12 environments we investigated—participants showed significant improvement in some conditions and virtually none in others. We examine this variation as a function of the consistency of feedback and entropy, finding that learning depends heavily on these characteristics of the decision environment.

Key Words: Change-point detection; regime shift; learning; over-reaction; under-reaction; Bayesian updating; probability judgment.

[1] University of California, Riverside, A. Gary Anderson Graduate School of Management, Riverside, CA 92521, ye.li@ucr.edu, (951) 827-3694. Corresponding author.
[2] University of Pennsylvania, Wharton School of Management, Philadelphia, PA 19104, cadem@wharton.upenn.edu, (215) 746-5884.
[3] University of Chicago, Booth School of Business, 5807 S. Woodlawn Ave., Chicago, IL 60637, wu@chicagobooth.edu, (773) 834-0519.

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 under-reaction 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 under-reacting (failing to realize change has occurred) against the costs of over-reacting (believing change has occurred when in fact it has not). For example, an investor needs to recognize when cyclical 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. Under-reacting to signals of change means foregoing the chance to buy stocks at their lowest prices, whereas over-reacting entails 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. Massey and Wu’s studies revealed a behavioral pattern consistent with system neglect: Under-reaction was most common in unstable systems with precise signals, whereas over-reaction was most prevalent in stable systems with noisy signals. Kremer, Moritz, and Siemsen (2011) corroborated and extended their work, finding evidence of 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 (e.g., Coursey, Hovis, & Schulze, 1987; List, 2003). In Massey and Wu (2005a; hereafter MW), for example, the diagnosticity and transition probability of the system changed after each of the 18 trials. On the one hand, this design enhanced the salience of the system variables, increasing the likelihood that participants would give them sufficient attention. On the other hand, this continuously-changing design may have hindered participants' ability to appropriately adjust to these dimensions by minimizing their opportunity to learn about a particular system and therefore raises questions about the robustness of the system-neglect hypothesis. Does the pattern of over- and under-reaction observed in MW persist in the face of learning? Or is the system-neglect phenomenon only observed in those inexperienced with the task and therefore less applicable to real-world settings where people have sufficient opportunities to learn? Finally, what system characteristics moderate the degree of learning? The present paper aims to fill these gaps in our understanding.

The rest of the paper is organized as follows. We begin by describing the statistical process used in our experiment and review the system-neglect hypothesis. Next, we present the experimental design and results for four different performance measures, in particular examining differences in learning between conditions. Then, we propose some psychological variables that seem to explain the differences in learning across environments. We conclude by discussing open questions and future directions.

BACKGROUND AND THEORY

In this section, we introduce the design of our experiment and review the system-neglect hypothesis predictions for detecting changes in this statistical process.

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*. The systems are dynamic in the sense that they can generate signals using two different sets of probabilities, each of which we call a *regime*. A system is characterized by two system variables: *diagnosticity*, or the informative of the signals it generates, and *transition* *probability*, or the likelihood of the system switching to the second regime.

Experimental Paradigm

Our experimental paradigm largely mirrors MW's. 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 the first period before any signal is drawn). If the system switches to the blue regime, it does not switch back. That is, the blue regime is an absorbing state.[4]

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 $p_B < .5$. Put differently, a red signal is more suggestive of a red regime, and a blue signal is more suggestive of a blue regime. 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 and 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 that governed each period of that trial.

Optimal responses to the task required application of Bayes' Rule. Let $B_i = 1$ ($B_i = 0$) indicate that the system is in the blue (red) regime in period i, and let $b_i = 1$ ($b_i = 0$) indicate that a blue (red) signal is observed in that period. If $H_i = (b_1, ..., b_i)$ is the history of signals through period i, the Bayesian posterior odds of a change to the blue regime after observing a history of

[4] Although having an absorbing state is a simplification, this is a standard assumption in this literature. An absorbing state process is also a reasonable way to model many of the low-frequency real-world examples we discuss above. To the investor who hopes to cash out before a downturn, it does not matter if a bear market would eventually return to its bull status—for the relevant decision timeframe, the bear market might as well be an absorbing state.

signals H_i is:

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

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 experience.

The System-Neglect Hypothesis

The system-neglect hypothesis posits that people are more sensitive to signals than to system variables. This hypothesis extends work by Griffin and Tversky (1992) that proposes that people 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 over-confidence when strength is high and weight is low, and under-confidence 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 over-react to signals of change in stable systems with noisy signals, and are more likely to under-react 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 under-reaction or only over-reaction.

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 $P(B_2 \mid H_2) = .17$ when d = 1.5 and q = .05 (*i.e.*, a noisy and stable system), but $P(B_2 \mid H_2) = .92$ when d = 9 and q = .20 (*i.e.*, a precise and unstable system). If individuals give approximately the same response across all four conditions (for example, with a posterior probability of .60), they will over-react when d = 1.5 and q = .05 and under-react 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 EXPERIMENT

Methods

We recruited 240 University of Chicago participants for a task advertised as a "probability estimation task." Each participant was assigned to one of the 12 experimental conditions, constructed by crossing three diagnosticity levels (d = 1.5, 3, and 9) with four transition probability levels (q = .02, .05, .10, and .20). These 12 systems were the same ones used in MW. The most important deviation from MW is switching to a between-subjects design: each participant only saw one set of system variables for 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 the Appendix.

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.[5]

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 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 and 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 by that period. 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 20 trials for actual pay. At the end of each paid trial, participants received

[5] Massey and Wu (2005a) also used a quadratic scoring scheme, although participants in that study could make or lose as much as 10 cents in each period.

feedback about which bin generated each ball, earnings for that trial, and cumulative earnings in the experiment.

Results

In this section, we summarize basic results for performance and learning. We look at two measures of performance: (1) earnings and (2) mean absolute difference between empirical and Bayesian judgments. To test the system-neglect hypothesis, we then consider measures of reaction to indications of change. Finally, we estimate a Quasi-Bayesian model to provide a formal test of the system-neglect hypothesis. In all cases, we investigate whether participants learn with experience over the course of the experiment, as well as how these measures vary across experimental conditions.

Earnings. Recall that we paid participants on the square of the difference between their subjective probability of having changed to the blue regime and the actual regime that period (1 if blue, 0 if red). The mean of this absolute difference between their predictions and the actual regime was .149 (median = .072, sd = .191), generating mean earnings of $11.05 total over 200 responses (range of $2.03 to $15.19 across participants). As a standard for comparison, an optimal Bayesian agent would have earned $12.95. Table 1 presents the mean empirical and Bayesian earnings for each condition and overall. Participants deviated from Bayesian earnings by $1.90 on average (range of $-0.27 to $12.30 across participants), earning 14.6% less than a Bayesian agent would earn.

To investigate whether earnings increased with experience, we first examined how earnings changed over the course of the 20 trials by dividing them into five quintiles of 4 trials each (results are virtually identical if we use quarters consisting of 5 trials). Table 1 lists the earnings by quintile, multiplied by five to be comparable to the overall earnings. Overall, the average earnings monotonically increased for each quintile, from \$10.73 in the first quintile to \$11.33 in the final quintile, and from 84.2% of Bayesian earnings to 86.6%. There was evidence of learning in most of the 12 conditions, with absolute earnings increasing from the first to last quintile in 8 of the 12 conditions and earnings as a proportion of Bayesian earnings increasing in 9 of the 12 conditions.

We conducted a more rigorous test of learning by running linear regressions of earnings on trial order for the 400 participant-trial observations per condition (20 participants per condition × 20 trials per participant), accounting for participant-level random-effects. The coefficients on trial are found at the bottom of Table 1, with positive numbers corresponding to improving performance. Although we assumed linear learning for simplicity, we will discuss nonlinear learning patterns below. Coefficients in 10 of the 12 conditions were positive and significant at the .05 level in the d = 9, q = .10 condition, and at the .10 level in 4 other conditions. We conducted the same analysis separately for each of the 240 participants, finding positive coefficients for 139 (58%) of the 240 participants overall (p = .017, two-tailed binomial test). The percentages of participants who had positive regression coefficients in each condition are shown at the bottom of Table 1.

****************INSERT TABLE 1 HERE****************

Mean absolute deviation. As evident from Table 1, earnings are a potentially noisy measure of performance, since the signals that participants respond to can be unrepresentative of the underlying regimes that determine their earnings. Indeed, Bayesian judgments can produce lower earnings than non-Bayesian judgments for small, unrepresentative sequences. We therefore 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 for MADs by condition and quintile.

Overall, MADs decreased over the five quintiles, though the pattern was not uniform across conditions. As before, we regressed MADs by trial for each condition, accounting for participant random-effects (see Table 2). The coefficients were negative (suggesting learning) for 11 of the 12 conditions and significant at the .05 level in 5 of the conditions, and at the .10 level in the $d = 3$, $q = .10$ condition. Separate regressions for each participant showed a similar but slightly stronger pattern to the analysis in Table 1, with 142 (or 59%) of the 240 participants having negative coefficients ($p < .01$, sign test). Overall, we found somewhat more pronounced evidence of learning for this less ambiguous measure of performance.

****************INSERT TABLE 2 HERE****************

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. Therefore, to test for system neglect, we compared participants' reactions (*i.e.*, the change in probability judgments from period $i-1$ to period i) to the Bayesian reaction using the participant's probability judgment in the previous period as the "prior" (see MW for more details). Importantly, this approach focuses only on reactions, granting participants their priors regardless of accuracy, and evaluating only how their judgments react to new information. We define errors in reaction as the difference between Bayesian and empirical reactions, with under-reaction indicating an empirical reaction that is less positive than the Bayesian reaction, and over-reaction indicating the opposite. Recall that the system-neglect hypothesis predicts a greater tendency to under-react in more precise, less stable conditions, and to over-react in noisier, more stable 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 and Wu, 2005a). As predicted by the system-neglect hypothesis, and replicating MW, the greatest under-reaction occurred in the southeast-most cells ($d = 9$ with $q = .05$, $q = .10$, and $q = .20$, and $d = 3$ with $q = .20$), while the greatest over-reaction occurred in the northwest-most cells ($d = 1.5$ with $q = .02$ and $q = .05$ and $d = 3$ with $q = .02$). For 41 of the 48 pairwise comparisons between conditions, under-reaction increased monotonically with diagnosticity and transition probability. For comparison, Figure 1 also plots the reactions from Massey and Wu (2005a). The pattern of system neglect is somewhat more pronounced in that study relative to the current investigation.

****************INSERT FIGURE 1 HERE****************

Figure 2 plots the same errors in reactions to blue signals by quintile. Note that the degree of system neglect is most pronounced in the first quintile but remains significant in the remaining four quintiles. For example, 39 of the 48 pairwise comparisons for the last quintile are in the direction of the system neglect hypothesis. Figure 2 also suggests that the learning that does take place occurs mostly in the highly precise conditions.

****************INSERT FIGURE 2 HERE****************

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 is an artifact of 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, α (sensitivity to transition probability) and β (sensitivity to diagnosticity):

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

Here, p_i^e denotes the empirical estimate at period *i,* α_m weights the transition probability for condition *m* = .02, .05, .10, and .20, and β_n weights diagnosticity for condition *n* = 1.5, 3, or 9, and. 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). System neglect requires that α and β vary by condition as follows: $\beta_{1.5} > \beta_3 > \beta_9$ and $\alpha_{.02} > \alpha_{.05} > \alpha_{.10} > \alpha_{.20}$ (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). We obtain similar results when we estimate parameters for each individual and average these estimates by condition. Figure 3 shows the estimates of α_m and β_n as well as the parameter pattern for complete system neglect. Here, lower levels imply greater conservatism in the sense of Edwards (1968). The estimated parameters were all ordered as predicted by system neglect, with pairwise differences significant at $p < .0001$ (*t*-statistics > 4.5), except for between β_3 and β_9 ($t = 2.40$, $p = .017$), and between $\alpha_{.10}$ and $\alpha_{.20}$ ($t = 1.39$, $p = .16$). In addition, there was a mixed pattern of conservatism and radicalism in both parameters, consistent with the pattern of over- and under-reaction depicted in Figure 1. Importantly, whereas the slopes were much steeper than 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.

****************INSERT FIGURE 3 HERE****************

****************INSERT FIGURE 4 HERE****************

To investigate whether system neglect decreased with experienced, we estimated α_m and β_n as before, adding dummies for each quintile. Figure 4 shows these estimates for the first and last quintile. The β parameter estimates for both moderately (d = 3) and highly precise (d = 9) conditions converged toward the Bayesian standard of 1 (ts = 1.58 and 4.18, respectively, p = .12 and p < .0001) but were still less than 1 (ts = -2.23 and -2.63, respectively, p < .05 and p < .01). In contrast, participants started the study assigning appropriate weight to noisy signals ($\beta_{1.5,Q1}$ = 1.03, t = 0.09, *ns*) and there was virtually no change by the end of the study (t = 0.11, *ns*).

All α parameter estimates converged toward the Bayesian standard, but the improvement was significant only for the q = .20 conditions (t = 2.62, p < .01), and insubstantial for more stable conditions ($t_{.02}$ = 0.41, $t_{.05}$ = 0.83, $t_{.10}$ = 0.17, *ns*). The α parameter estimates in the last quintile remained less than 1 for the more unstable conditions ($t_{.10}$ = -5.70 and $t_{.20}$ = -7.17, ps < .0001), but was not significantly different from 1 for the more stable conditions ($t_{.02}$ = 1.21 and $t_{.05}$ = -0.58, *ns*).

These results are again quite consistent with the earlier analyses on earnings and MADs, and additionally reveal that most of the learning corresponds to less conservative responses to precise signals and when there is a high base-rate of change.

Explaining Differences in Learning across conditions

So far, we have seen that although system neglect is not completely eliminated, average performance does improve with experience. We have also demonstrated that this improvement differs substantially across conditions. In this section, we try to understand why experience leads to learning in some domains and not in others.

Figure 5 plots performance—as measured by mean absolute deviation from Bayesian judgments—across conditions and reveals a great deal of variation in initial performance, final performance, total learning, and shape of the learning over the 20 trials.[6] Importantly, initial performance (as measured by MAD in the first quintile) was orthogonal to learning ($r = .04$, *ns*). For example, although initial performance was roughly equal in all three $q = .05$ conditions, the $d = 9$ condition showed significant learning ($t = -2.11$, $p < .05$), the $d = 3$ condition showed no learning ($t = -0.48$, *ns*), and the $d = 1.5$ condition actually got worse with experience, although this decrement was not significant ($t = 0.64$, *ns*). Similar differences can be found for comparisons between the $d = 3$, $q = .02$ and $d = 9$, $q = .02$ conditions, and between the $d = 1.5$, $q = .20$ and the $d = 3$, $q = .20$ conditions.

****************INSERT FIGURE 5 HERE****************

Psychological determinants of learning. 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). To understand the variation in learning across conditions, we consider basic psychological determinants of learning and consider how these concepts apply to the change-detection paradigm. At its most basic level, learning from past behavior requires informative feedback about the consequences of that behavior.[7]

---

[6] These curves were generated using LOWESS (locally-weighted scatterplot smoothing), a non-parametric method for plotting a smoothed curve by computing a localized regression for the nearest 0.5*n* neighbors for each data point.

[7] Note that we have chosen to investigate the environment explicitly, rather than the learning process itself. This approach differs from other research on learning in stochastic environments that directly examines the reinforcement learning process (Bush & Mosteller, 1953; Camerer & Ho, 1999; Cross, 1983; Erev & Roth, 1998). We adopt our approach for two reasons. First, just as we focus on system variables to understand when system neglect leads to over- and under-reaction, 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

In our experiment, participants received explicit feedback about when the regime actually shifted and therefore what series of probability judgments would have maximized payment for that trial. Although this feedback is unambiguous, it was provided only at the end of each trial and therefore was not entirely informative or generalizable (*i.e.*, knowing the actual regime in each period does not provide the optimal reactions to each signal). On the other hand, participants also received implicit feedback about their judgments from subsequent draws. Indeed, although ambiguous, this period-level feedback is more immediate and extensive than the trial-level feedback. We therefore explore two aspects of period-level feedback that may vary across conditions: (1) how consistent the feedback is, and (2) the informativeness of feedback for a particular type of judgment.

We first consider feedback consistency. The central task in change detection is monitoring and responding to the appearance of a blue signal, which indicates the possibility of change. Although the informativeness of the signal is determined by the system's diagnosticity, the information provided by a blue signal can be reinforced or undermined by whether the next signal is consistent (*i.e.*, also a blue signal, strongly suggesting a regime shift) or not (*i.e.*, a red signal, suggesting a false alarm). That is, the signal following each probability judgment provides feedback about the accuracy of that prior judgment. Even though this feedback is noisy, it is immediate, clear, and exceedingly vivid, all qualities that greatly facilitate learning (Hogarth, 2001; Maddox, Ashby, & Bohil, 2003; Nisbett & Ross, 1980). We therefore define the *inconsistency* of an environment as the proportion of times that a new signal of change (*i.e.*, a blue signal not immediately preceded by another blue signal) is followed immediately by a

---

certain responses in various scenarios, learning would move α and β toward more optimal parameter values. Although such an approach seems reasonable, it would be quite a departure from previous research.

signal of non-change (*i.e.*, a red signal). A higher proportion of inconsistent feedback should make it more difficult to learn.

A second fundamental aspect of feedback is its information value. A convenient way to measure information value is entropy (Shannon, 1948), $H(x) = -\sum p(x_i)\ln\left(p(x_i)\right)$. In our setting, we consider the entropy of the Bayesian posterior probabilities of change (indicating the likelihoods of regime shift). Figure 6 shows how dramatically the Bayesian posterior distributions vary across our experimental conditions, from those with virtually all optimal judgments at or near 0 or 1, to those with almost all optimal judgments *between* 0 and 1.

****************INSERT FIGURE 6 HERE****************

The entropy of these distributions depends on how uniform versus "spiky" it is. A uniform distribution has maximal entropy—and minimal information value—because it conveys the least information possible about the true state. In contrast, a point distribution has zero entropy, which is the maximum information value, as it conveys with perfect certainty the true state. In our task, noiseless conditions, as well as very stable or unstable conditions, have low entropy with nearly binary Bayesian posteriors clustered near 0 and 1. On the other hand, conditions with noisy signals and intermediate transition probabilities have Bayesian posteriors more uniformly distributed on the intermediate probabilities, and thus high entropy.

Entropy may be negatively related to learning for a few potentially related reasons. Most importantly, it is much easier to respond to signals of change in a binary fashion than to distinguish between more nuanced levels of probability. Feedback is clear and easy to encode when it reinforces binary responses of 0 or 1, whereas feedback that reinforces intermediate probabilities has lower informational value and is thus harder to encode. This difficulty of encoding intermediate probabilities is consistent with research on the probability weighting function, indicating that individuals do not discriminate intermediate probabilities as well as they do more extreme ones (Wu & Gonzalez, 1996). Research on preference reversals also suggests that our inherently binary stimuli might induce more binary responses (Fischer & Hawkins, 1993; Tversky, Sattath, & Slovic, 1988).

Another possible reason for entropy to detract from learning is that systems with higher entropy provide more *exacting* feedback (Hogarth, McKenzie, Gibbs, & Marquis, 1991). Due to shape of the payoff function, intermediate probability judgments are punished (and rewarded) less than extreme ones. In the extreme case, a response of 0.5 in our study is always rewarded with 4 cents regardless of the actual regime, whereas a response of 0 or 1 is rewarded or punished 8 cents depending on whether the response matches the actual regime. Hence, in systems with low entropy—those in which most of the optimal responses are near 0 or 1—there is a much stronger incentive to be accurate.

Feedback inconsistency, entropy, and learning

Table 3 shows the proportion of inconsistent feedback and optimal response entropy across the 12 experimental conditions. For our range of system parameters, entropy monotonically decreases with signal diagnosticity but has a curvilinear relationship with system stability: Very stable (q = .02) and very unstable (q = .20) systems have low entropy whereas systems with intermediate transition probabilities have higher entropy. Inconsistency of feedback decreases with diagnosticity, and somewhat more so for more unstable systems.

**************INSERT TABLE 3 HERE**************

Although we only have 12 conditions as data points, we cautiously explore how these psychological variables related to learning. Also, although Figure 5 shows that learning is clearly nonlinear and sometimes even non-monotonic, it also shows that simple linear regressions capture most of the learning in each condition. We therefore use the linear coefficients of MAD regressed on trial (see Table 2) to index learning as a simple approximation of the overall degree of learning. Analyses using alternative measures, such as learning between the first quintile to the best quintile (to account for later performance decrements due to potential fatigue), from the first quintile to the second quintile (as a measure of early learning), and from the third quintile to the last quintile (as a measure of continued learning) yield qualitatively similar results.

Table 4 shows the results of regressions of the linear learning coefficient on each of the psychological variables, as well as d, q, and their interaction. Due to the high correlations between feedback consistency and entropy ($r = .79$, $p < .01$), we do not present the results of models that include both psychological variables. In addition, we present a second set of regressions controlling for initial performance as measured by MAD in the first quintile. Doing so helps reveal learning that is obscured by a floor effect on learning: It seems unlikely for MAD to decrease below the .05 level. Controlling for initial performance therefore helps explain the lack of learning in conditions that start too well to have much learning.

****************INSERT TABLE 4 HERE****************

We can draw a few interesting inferences from these regression results. First, the system parameters do not explain learning at all, whether controlling for initial performance or not. Second, initial performance has a negative effect on learning, confirming our intuition that conditions in which participants had low MADs early on provided minimal opportunity for learning. Finally, feedback inconsistency was a strong deterrent of learning, explaining 40% of the variance in learning on its own. Entropy was also negatively related to learning but not significantly so due to our small sample size. However, controlling for initial performance strengthened the relationship between entropy and learning to significance, and revealed a perhaps even stronger relationship than the between consistent feedback and learning.

DISCUSSION

We examined learning across 12 conditions varying on signal diagnosticity and system stability, using a number of measures of performances, including earnings, absolute deviations from Bayesian estimates, or reactions. In all cases, learning was more prevalent in highly precise conditions and the moderately precise, more unstable conditions. We also found that system neglect was partially attenuated by 20 trials of experience, especially for moderately and highly precise conditions, as well as for highly unstable conditions.

In order to better understand this variation in learning across conditions, we more directly examined how two psychological characteristics of the learning environments corresponding to the consistency and informational value of feedback predicted learning. We found more learning in conditions with lower entropy—and thus more binary optimal responses—and more consistent feedback—less likelihood of getting contradictory period-level feedback.

Although these results are suggestive, we caution that they capture the relationship between the psychological variables and learning on only 12 conditions—with only three levels of diagnosticity and four levels of transition probability—and only 20 sequences per condition.[8] Although we can simulate the psychological variables for a larger number of series and a wider range of system variables, we cannot simulate participants' performance and learning. It is possible that these results are different for a wider range of system variables. Perhaps more importantly, we did not directly manipulate these variables and therefore cannot establish their causality. We discuss this issue in more detail in the next section.

[8] We do note that a remarkably similar pattern of results hold for an earlier pilot study with only six conditions (the same three levels of diagnosticity but only q = .05 and .10). That is, feedback inconsistency was the best predictor of learning.

Psychological determinants of learning

On some level, entropy and trickiness are necessarily complex, nonlinear transformations of some combination of diagnosticity and transition probability. It is therefore possible that there are even better transformations that we have not uncovered. Although these psychological variables may be related to the system parameters, other elements of a learning environment can also affect entropy and how inconsistent the feedback is, independently of the system parameters. Future research could therefore isolate the effect of the consistency of feedback or its informational value by directly manipulating them while holding the system parameters constant.

For example, we can change the consistency of period-level feedback by minimizing the number of false signals in more tricky environments, perhaps by changing to five periods of two signals each, akin to providing feedback every few trials as in Lurie and Swaminathan (2009). Or even more directly, we can ask for judgments of the previous periods' urn.

Similarly, we can manipulate task design to directly manipulate entropy. Recall that entropy may deter learning for one or both of two possible reasons. First, higher entropy environments require learning to respond with a continuous range of probabilities, which is harder than learning to respond in a binary fashion, both for informational and perceptual reasons. Second, the payment function is relatively flat for intermediate probabilities and therefore feedback is less clear. Either one of these factors could drive higher entropy environments to be harder to learn in. More critically, experimentally manipulating these factors could help to isolate their relative contribution. For example, instead of paying based on deviation from the actual regime, we can pay based on deviation from the Bayesian judgment (possibly with an error term) to increase the sharpness of the payment curve for intermediate probabilities. In addition, we could compare our results to stationary probabilistic judgment tasks (such as the "balls and urns" task of Edwards, 1968). In these settings, it is possible to have low entropy with Bayesian posteriors clustered around 0.5 (instead of around 0 and 1, as in our task).

More generally, we draw attention to the usefulness of entropy for explaining judgment and learning. Entropy is a central foundation of the field of information theory, and has been used in theoretical models of learning (Lehrer & Smorodinsky, 2000). Although entropy has been shown to influence the learning of both pigeons (Young & Wasserman, 1997) and humans (Kvålseth, 1978), and more recently invoked to explain anomalies in behavioral finance and economics (Peng, 2005; Peng & Xiong, 2006; Sims, 2003), it is rarely the focus of psychological research. We believe entropy is potentially a rich construct that deserves more exploration in studies of judgment and learning.

Organizational decision-making

The influence of environmental factors on learning has direct implications for organizations. Obviously it would be helpful to increase the quality or quantity of feedback available to a decision maker. For example, instituting a waiting period before reacting to signals of change helps to reduce feedback inconsistency and therefore rates of reacting to a false alarm. Unfortunately firms often do not or cannot control the feedback available in their environment. However, they may able to improve decision-makers' *attention* to feedback, by 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 environments with inconsistent feedback.

Conclusion

Our paper establishes the robustness of system neglect in change-point detection and demonstrates the relationship between psychological characteristics of an environment and learning. In the end, we are somewhat sober about the ability of individuals to avoid systematic over- and under-reaction 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

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.

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.

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.

Brown, S. D., & Steyvers, M. (2009). Detecting and predicting changes. *Cognitive Psychology,* 58(1), 49-67.

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.

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.

Deming, W. E. (1975). On Probability as a Basis for Action. *Amer. Statistician,* 29(4), 146-152.

Edwards, W. (1968). Conservatism in human information processing. *Formal representation of human judgment,* pp. 17-52.

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. 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., & Lattin, J. M. (1993). Accounting for Heterogeneity and Nonstationarity in a Cross-Sectional Model of Consumer Purchase Behavior. *Marketing Sci.,* 12(3), 304-317.

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. *Research in organizational behavior,* pp. 1-37.

Hogarth, R. M. (2001). *Educating intuition*. Chicago, IL: University of Chicago Press.

Hogarth, R. M., McKenzie, C. R. M., Gibbs, B. J., & Marquis, M. A. (1991). Learning From Feedback: Exactingness 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(10), 1827-1843.

Kvålseth, T. O. (1978). The effect of task complexity on the human learning function. *The International Journal of Production Research,* 16(5), 427-435.

Lehrer, E., & Smorodinsky, R. (2000). Relative entropy in sequential decision problems. *Journal of Mathematical Economics,* 33(4), 425-439.

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.

Maddox, W. T., Ashby, F. G., & Bohil, C. J. (2003). Delayed feedback effects on rule-based and information-integration category learning. *Journal of Experimental Psychology-Learning Memory and Cognition,* 29(4), 650-662.

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"). http://faculty.chicagobooth.edu/george.wu/research/change_points_ecp.pdf.

McFadden, D. (1981). Econometric Models of Probabilistic Choice. *Structural Analysis of Discrete Data with Econometric Applications,* pp. 198-272.

Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*: Prentice-Hall Englewood Cliffs, NJ.

Peng, L. (2005). Learning with information capacity constraints. *Journal of Financial and Quantitative Analysis,* 40(2), 307-329.

Peng, L., & Xiong, W. (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics,* 80(3), 563-602.

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.

Shannon, C. E. (1948). The mathematical theory of communication. *Bell System Technical Journal,* 27, 379-423.

Shewhart, W. A. (1939). *Statistical Method from the Viewpoint of Quality Control*. Washington: Graduate School, Department of Agriculture.

Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics,* 50(3), 665-690.

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.

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 period, and 1 indicating a blue ball for that period. The period that the regime shifts to the blue regime, if ever, is noted in the "Shift" column. For example, sequence 4 of the $d = 1.5$, $q = .02$ condition experienced a shift to the blue regime in period 5, meaning that the first four draws were from the red regime and the fifth through tenth draws were from the blue regime.

****************INSERT TABLE A1 HERE****************

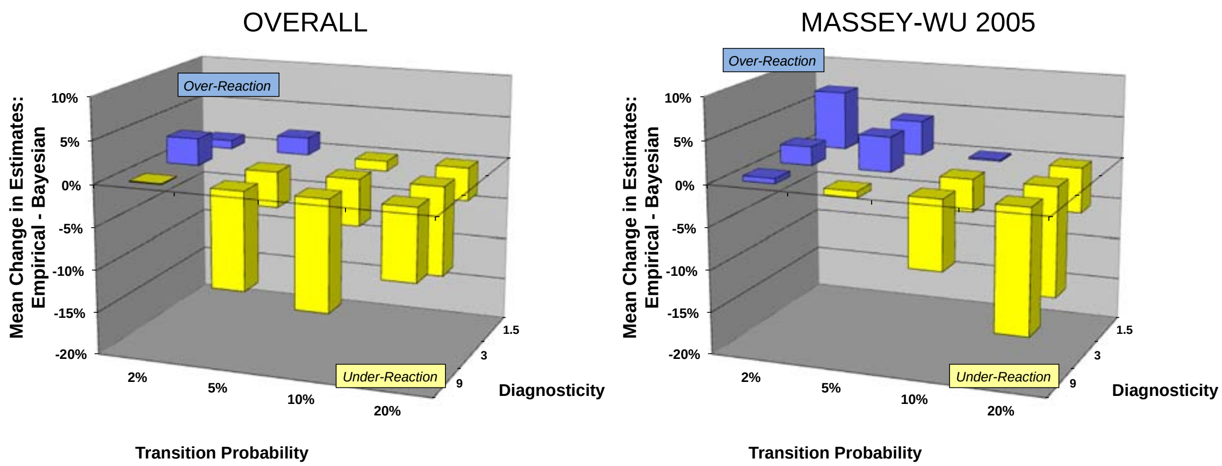


Figure 1: Over- and under-reaction to blue signals, by condition, as measured by the mean difference between the change in empirical probability judgments and Bayesian reaction. The left panel shows this measure for the current study. The right panels show the same measure of Massey and Wu (2005a).

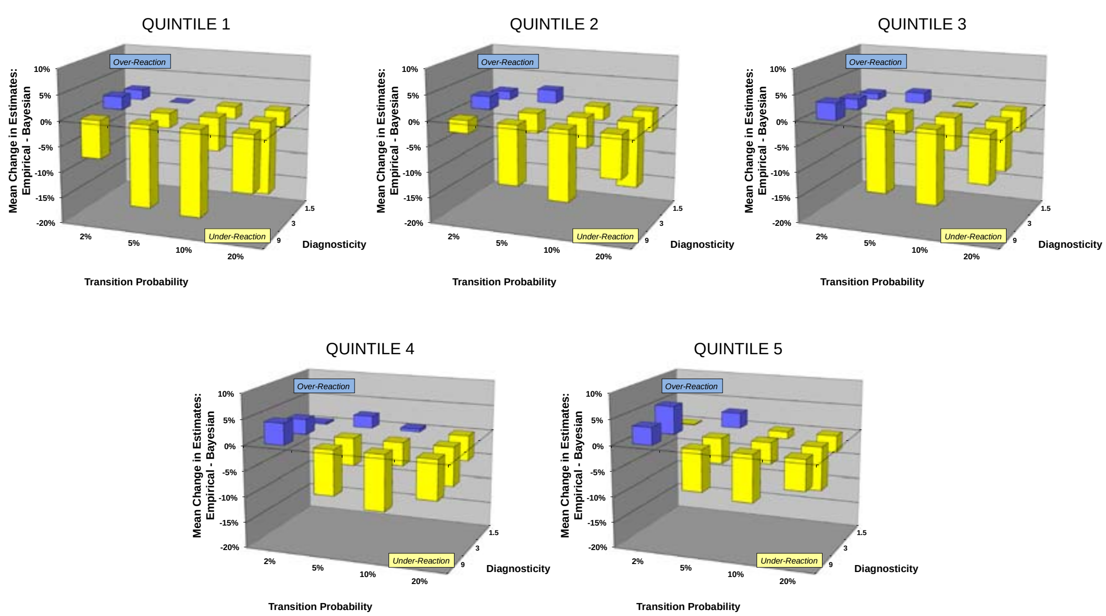


Figure 2: Over- and under-reaction, by condition and quintile, as measured in Figure 1.

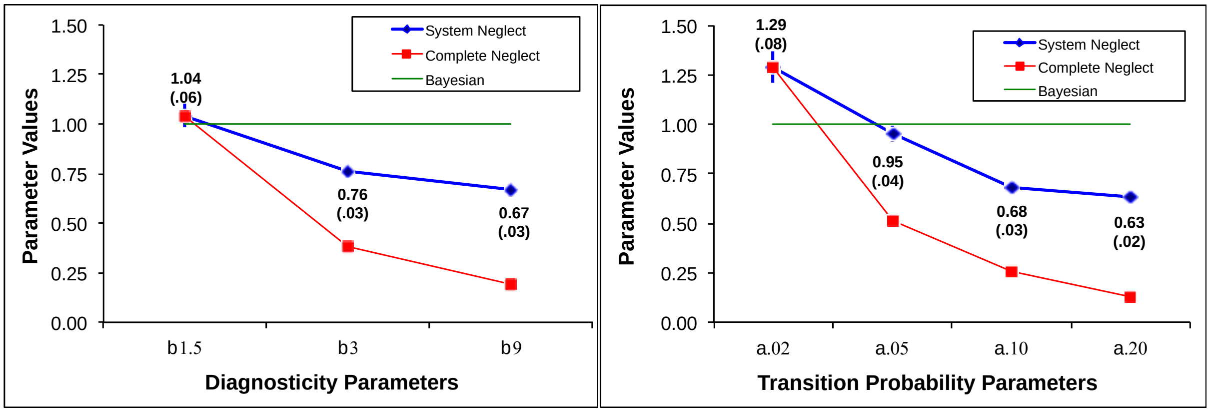
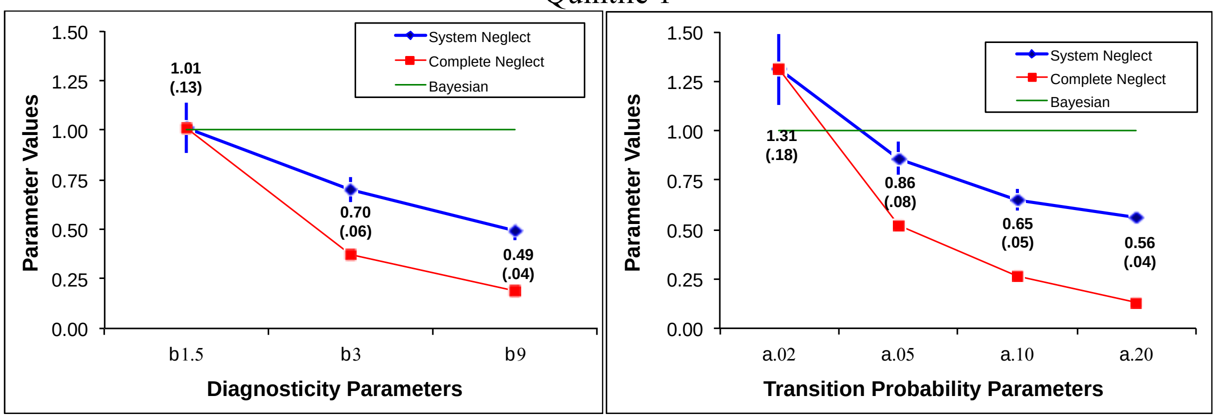


Figure 3. System neglect parameter estimates by condition, as fit to Quasi-Bayesian model.

Quintile 1



Quintile 5

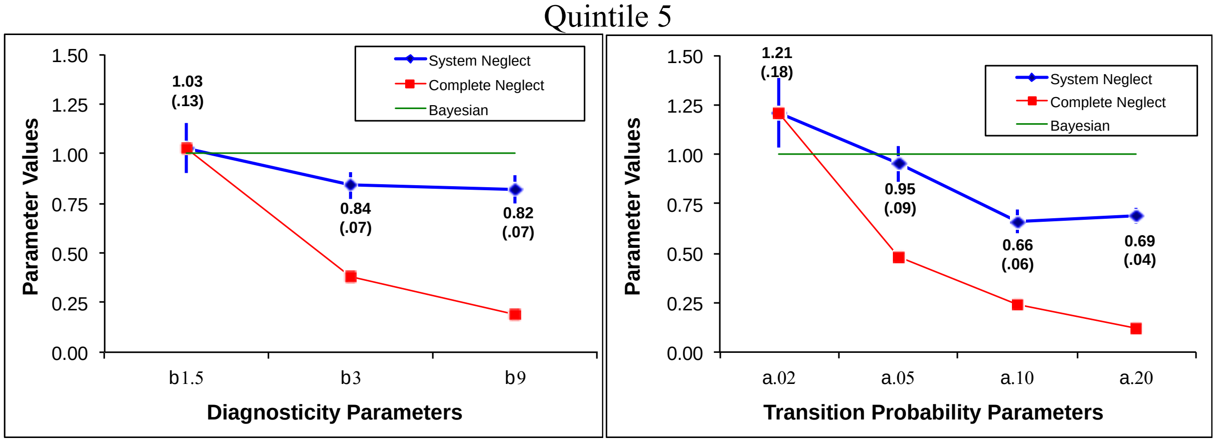


Figure 4. System neglect parameter estimates by condition for first and last quintile, as fit to Quasi-Bayesian model.

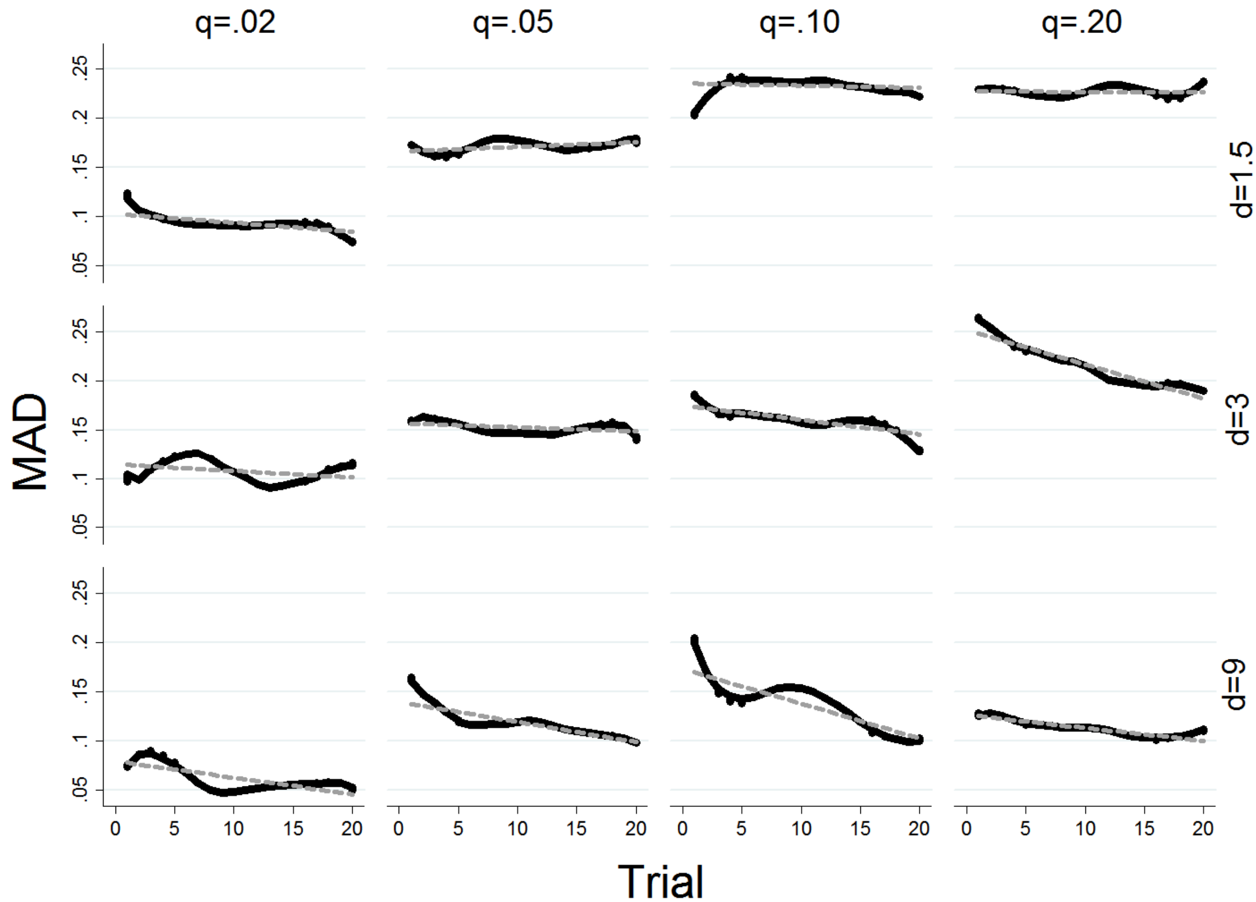


Figure 5. LOWESS (locally-weighted scatterplot smoothing, using a bandwidth of 0.5n) and linear regressions of mean absolute deviation from Bayesian judgment (MAD) on trial.

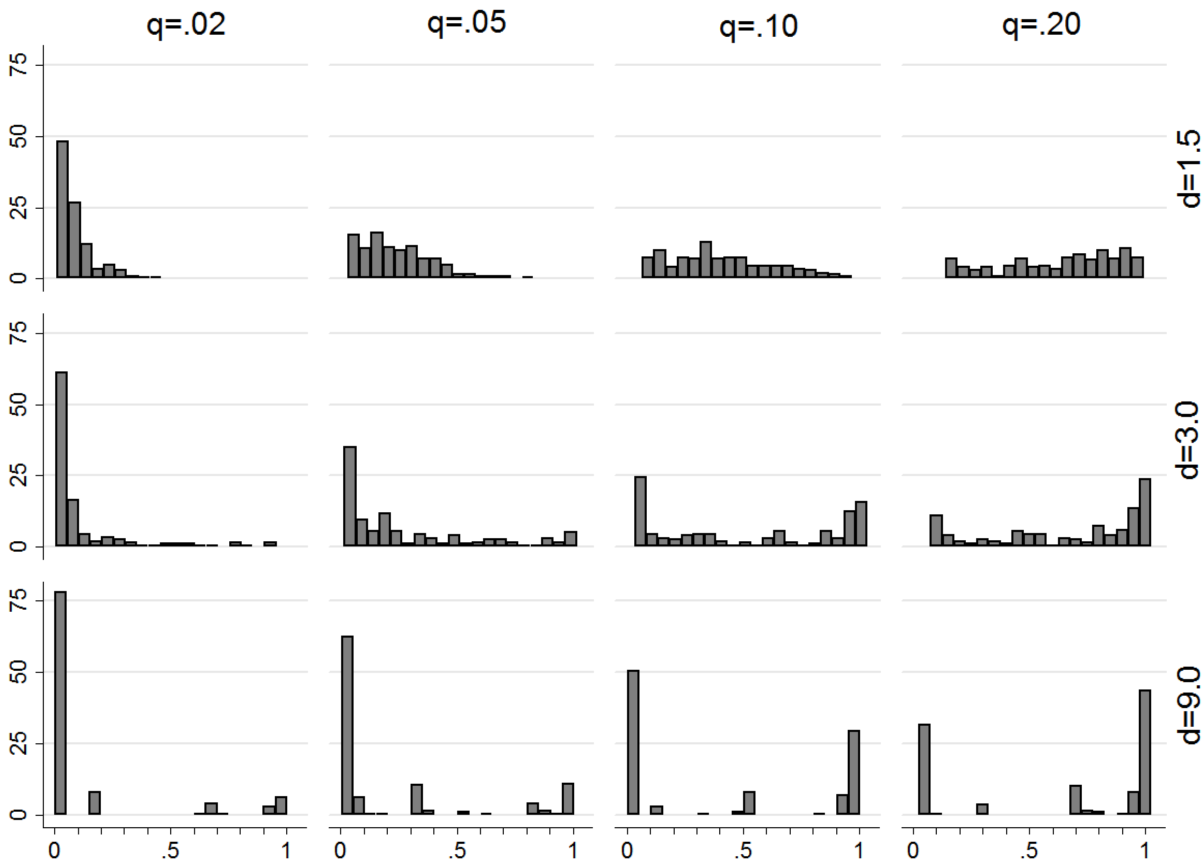


Figure 6. Distribution of Bayesian posteriors by condition, binned by .05 intervals.

	Condition												
d	1.5	1.5	1.5	1.5	3	3	3	3	9	9	9	9	
q	.02	.05	.10	.20	.02	.05	.10	.20	.02	.05	.10	.20	Overall
Empirical Earnings													
Earnings	14.35	9.69	7.88	6.59	12.36	10.92	10.55	8.84	14.15	12.30	12.08	12.96	11.06
Standard Deviation	0.88	1.71	1.32	1.24	1.39	0.84	2.50	2.69	1.61	1.32	0.57	1.75	2.82
Bayesian Earnings													
Earnings	14.92	11.47	10.21	9.85	13.49	12.73	12.44	12.20	15.07	14.05	14.70	14.28	12.95
Empirical/Bayesian	96.2%	84.5%	77.2%	66.9%	91.6%	85.8%	84.8%	72.5%	93.9%	87.5%	82.2%	90.8%	85.4%
Empirical Earnings by Quintile (scaled by 5)													
Quintile 1	14.18	8.74	8.23	6.87	12.93	10.70	9.81	7.69	13.10	11.63	11.86	12.98	10.73
Quintile 2	14.31	10.16	8.62	6.60	11.14	11.43	10.36	8.88	14.30	12.13	11.33	12.94	11.02
Quintile 3	14.35	9.57	7.85	6.48	12.50	10.38	10.82	9.11	14.28	12.54	11.75	13.01	11.05
Quintile 4	14.14	10.34	6.49	5.56	13.10	11.42	10.48	9.71	14.87	12.57	11.97	13.20	11.15
Quintile 5	14.76	9.64	8.23	7.43	12.12	10.68	11.26	8.82	14.18	12.61	13.49	12.69	11.33
Bayesian Earnings by Quintile (scaled by 5)													
Quintile 1	14.77	11.06	10.46	9.02	13.19	12.54	11.62	11.73	15.02	14.08	14.80	14.60	12.74
Quintile 2	15.02	11.09	10.36	9.52	13.31	12.71	12.18	12.32	15.13	13.91	14.70	14.19	12.87
Quintile 3	15.01	11.40	10.23	10.05	13.68	12.66	12.50	12.58	15.11	14.03	14.53	14.16	13.00
Quintile 4	14.92	11.78	10.08	10.38	13.86	12.69	12.74	12.46	15.05	14.17	14.54	14.25	13.08
Quintile 5	14.90	12.00	9.93	10.29	13.43	13.06	13.15	11.92	15.02	14.05	14.96	14.19	13.08
Empirical/Bayesian Earnings by Quintile													
Quintile 1	96.0%	79.0%	78.7%	76.2%	98.0%	85.3%	84.4%	65.5%	87.2%	82.6%	80.1%	88.9%	84.2%
Quintile 2	95.3%	91.6%	83.2%	69.3%	83.7%	89.9%	85.1%	72.1%	94.5%	87.2%	77.1%	91.2%	85.6%
Quintile 3	95.6%	83.9%	76.7%	64.5%	91.4%	82.0%	86.6%	72.4%	94.5%	89.4%	80.9%	91.9%	85.1%
Quintile 4	94.8%	87.8%	64.4%	53.6%	94.5%	90.0%	82.3%	77.9%	98.8%	88.7%	82.3%	92.6%	85.3%
Quintile 5	99.1%	80.3%	82.9%	72.2%	90.2%	81.8%	85.6%	74.0%	94.4%	89.8%	90.2%	89.4%	86.6%
Change in earnings by trial (scaled by 20)													
Regression coefficient	0.025	0.046	-0.058	-0.001	0.013	0.007	0.077	0.079	0.062	0.071	0.103	0.001	0.035
Standard error	0.036	0.065	0.05	0.050	0.050	0.047	0.042†	0.045†	0.033†	0.043†	0.047*	0.026	0.013**
% Positive Learning	45%	50%	45%	55%	60%	45%	75%	65%	65%	65%	80%	45%	58%

Table 1. Mean empirical and Bayesian earnings, as well as empirical earnings as a proportion of Bayesian earnings, by condition and quintile (set of 4 trials). Also shown are random-effects linear regression coefficients of earnings change by trial, and percentage of participants with improving earnings. Data by quintile and trial are re-scaled to facilitate comparison with overall earnings. † $p < .10$; * $p < .05$; ** $p < .01$

	Condition												
d	1.5	1.5	1.5	1.5	3	3	3	3	9	9	9	9	
q	.02	.05	.10	.20	.02	.05	.10	.20	.02	.05	.10	.20	Overall
Mean absolute deviation from Bayesian (MAD)													
Overall	0.094	0.171	0.233	0.227	0.107	0.152	0.160	0.215	0.062	0.118	0.136	0.112	0.149
Standard Deviation	0.076	0.102	0.112	0.115	0.114	0.108	0.107	0.140	0.091	0.119	0.166	0.116	0.127
Quintile 1	0.100	0.163	0.230	0.230	0.098	0.168	0.172	0.254	0.095	0.140	0.151	0.128	0.161
Quintile 2	0.096	0.173	0.236	0.222	0.136	0.143	0.168	0.226	0.052	0.110	0.153	0.108	0.152
Quintile 3	0.092	0.173	0.240	0.227	0.103	0.147	0.152	0.213	0.053	0.128	0.148	0.115	0.149
Quintile 4	0.094	0.174	0.230	0.233	0.087	0.143	0.162	0.188	0.049	0.108	0.131	0.106	0.142
Quintile 5	0.086	0.175	0.231	0.224	0.113	0.161	0.144	0.194	0.060	0.105	0.096	0.105	0.141
Change in MAD by trial													
Regression coefficient	-0.00092	0.00047	-0.00022	-0.00005	-0.00067	-0.00041	-0.00147	-0.00348	-0.00168	-0.00205	-0.00354	-0.00131	-0.0013
Standard error	0.00063	0.00073	0.00069	0.00074	0.00099	0.00085	0.00077+	0.00087***	0.00077*	0.00097*	0.00129**	0.00062*	0.00024***
% Positive Learning	60%	45%	55%	50%	60%	55%	70%	65%	55%	70%	70%	55%	59%

Table 2: Mean absolute deviations (MADs) between empirical judgments and Bayesian judgments, by quintile and overall. Also, random-effects linear regression coefficients of MAD change by trial and percentage of participants with improving MADs. † $p < .10$; * $p < .05$; ** $p < .01$;*** $p < .001$

		q = .02	q = .05	q = .10	q = .20
Entropy	d = 1.5	1.546	2.434	2.769	2.737
	d = 3	1.499	2.386	2.479	2.465
	d = 9	0.810	1.365	1.266	1.343
Feedback Inconsistency	d = 1.5	0.742	0.825	0.825	0.821
	d = 3	0.767	0.750	0.625	0.721
	d = 9	0.421	0.475	0.413	0.408

Table 3. Entropy (uniformity of the distribution of Bayesian judgments) and feedback inconsistency (proportion of new blue signals followed by a red signal) by condition.

	(1)	(2)	(3)	(4)	(5)	(6)
Initial Performance				-0.086	-0.101	-0.293
				[0.142]	[0.060]	[0.085]**
d	-0.003			-0.003		
	[0.002]			[0.002]		
q	-0.084			-0.008		
	[0.082]			[0.152]		
$d \times q$	0.009			0.002		
	[0.015]			[0.019]		
Feedback Inconsistency		0.046			0.061	
		[0.018]*			[0.019]*	
Entropy			0.008			0.027
			[0.005]			[0.007]**
Constant	0.005	-0.043	-0.027	0.014	-0.036	-0.018
	[0.009]	[0.012]**	[0.011]*	[0.019]	[0.012]*	[0.008]
R^2	0.42	0.40	0.17	0.45	0.54	0.64
Adjusted R^2	0.21	0.34	0.08	0.14	0.44	0.56

Table 4. Regressions of the linear learning coefficient on system variables and psychological variables, controlling for initial performance or not. More negative coefficients indicate greater learning.

			Period										
#	d	q	1	2	3	4	5	6	7	8	9	10	Shift
1	1.5	.02	1	1	0	0	0	0	0	1	0	0	-
2	1.5	.02	0	1	1	0	0	1	0	0	1	1	-
3	1.5	.02	0	0	0	0	0	0	0	0	0	0	-
4	1.5	.02	1	0	0	0	1	1	1	1	0	0	5
5	1.5	.02	1	0	1	1	0	0	0	1	0	0	-
6	1.5	.02	0	0	0	1	0	1	1	1	0	1	-
7	1.5	.02	0	0	1	0	0	0	1	1	0	0	-
8	1.5	.02	1	1	0	0	1	1	0	1	1	1	-
9	1.5	.02	0	0	0	0	0	0	0	0	0	1	-
10	1.5	.02	0	0	1	0	0	0	0	0	0	0	-
11	1.5	.02	1	0	0	0	1	1	0	0	0	1	-
12	1.5	.02	0	0	0	1	1	0	0	1	1	0	-
13	1.5	.02	1	1	1	0	0	0	1	1	1	0	-
14	1.5	.02	0	1	0	1	1	0	0	1	0	0	-
15	1.5	.02	1	0	1	1	1	1	0	0	1	1	-
16	1.5	.02	1	1	0	0	1	0	0	0	0	0	-
17	1.5	.02	0	0	1	0	0	1	0	1	1	1	10
18	1.5	.02	0	1	0	0	1	0	0	0	0	1	-
19	1.5	.02	0	0	0	1	0	0	1	0	0	1	-
20	1.5	.02	0	0	0	0	0	1	1	0	0	1	-
1	1.5	.05	1	1	1	0	0	1	1	0	1	1	2
2	1.5	.05	1	0	0	0	1	0	0	0	1	1	-
3	1.5	.05	0	1	1	0	0	1	0	1	0	0	-
4	1.5	.05	0	0	1	1	1	0	1	0	0	0	5
5	1.5	.05	1	0	1	1	1	1	0	1	1	1	9
6	1.5	.05	1	1	1	0	1	0	0	1	0	1	-
7	1.5	.05	0	0	1	0	0	1	0	1	1	0	-
8	1.5	.05	0	0	1	1	0	1	0	1	1	0	-
9	1.5	.05	0	0	0	1	0	1	0	0	0	0	4
10	1.5	.05	1	1	0	1	0	1	0	0	1	0	-
11	1.5	.05	0	0	1	0	1	0	0	0	1	1	-
12	1.5	.05	0	1	1	0	0	1	0	1	1	0	-
13	1.5	.05	0	1	1	1	0	1	1	1	0	1	8
14	1.5	.05	0	0	1	1	1	0	1	0	0	1	-
15	1.5	.05	1	0	0	0	0	0	1	0	1	1	-
16	1.5	.05	0	1	1	0	0	1	1	0	0	1	3
17	1.5	.05	1	1	0	0	0	0	0	0	0	0	-
18	1.5	.05	0	1	0	1	0	0	0	1	0	1	-
19	1.5	.05	0	1	0	1	1	0	0	0	0	1	-
20	1.5	.05	0	0	1	1	0	1	1	0	1	0	4

			Period										
#	d	q	1	2	3	4	5	6	7	8	9	#	Shift
1	3	.02	0	0	1	0	1	0	1	0	0	1	-
2	3	.02	0	1	1	1	0	0	1	1	1	1	2
3	3	.02	0	1	1	0	0	0	0	0	0	0	-
4	3	.02	0	0	0	0	0	1	0	0	0	0	-
5	3	.02	0	0	0	1	0	0	0	0	1	1	-
6	3	.02	0	0	0	0	0	1	0	0	0	0	-
7	3	.02	0	1	0	0	0	0	0	0	0	0	-
8	3	.02	0	0	0	0	0	0	0	1	0	1	-
9	3	.02	0	0	1	1	1	1	1	0	1	0	2
10	3	.02	1	1	0	0	0	0	0	1	1	1	-
11	3	.02	0	0	1	1	0	0	0	0	0	0	-
12	3	.02	0	1	1	1	0	0	1	1	0	1	2
13	3	.02	0	0	0	0	0	1	0	0	1	1	8
14	3	.02	0	0	0	0	0	0	0	1	0	0	-
15	3	.02	0	0	0	0	0	0	0	0	0	1	-
16	3	.02	0	1	0	0	0	1	0	0	0	1	-
17	3	.02	0	1	1	0	0	0	0	0	0	0	-
18	3	.02	0	1	0	0	1	0	1	0	0	0	-
19	3	.02	1	0	0	1	0	1	1	1	0	0	-
20	3	.02	1	0	0	1	0	0	1	0	0	1	10
1	3	.05	1	1	1	0	0	0	0	0	0	1	-
2	3	.05	0	0	0	1	1	1	1	1	1	1	4
3	3	.05	1	0	0	0	0	1	1	1	1	1	4
4	3	.05	0	0	1	1	0	0	1	0	0	0	-
5	3	.05	0	0	0	0	0	1	0	0	1	0	-
6	3	.05	0	1	0	0	1	0	1	1	0	0	-
7	3	.05	0	0	0	0	0	0	1	1	0	0	-
8	3	.05	0	0	0	1	0	0	1	0	1	0	-
9	3	.05	1	1	1	1	1	1	0	0	0	0	1
10	3	.05	0	1	0	0	0	0	1	0	1	1	6
11	3	.05	0	0	0	1	0	0	0	0	0	0	-
12	3	.05	1	1	1	1	1	1	1	1	0	1	1
13	3	.05	0	0	0	0	0	0	0	0	1	0	-
14	3	.05	0	0	0	0	0	0	0	0	1	1	-
15	3	.05	1	0	1	1	0	0	0	0	0	1	-
16	3	.05	0	0	1	1	0	1	1	0	0	0	-
17	3	.05	0	1	1	0	0	1	1	0	0	1	-
18	3	.05	0	0	0	1	1	0	0	0	1	0	-
19	3	.05	0	1	0	1	1	0	1	1	0	1	4
20	3	.05	0	1	0	0	1	0	1	1	0	1	6

			Period										
#	d	q	1	2	3	4	5	6	7	8	9	#	Shift
1	9	.02	0	0	0	0	0	0	0	0	0	0	-
2	9	.02	0	0	0	0	0	0	0	0	0	1	-
3	9	.02	0	0	0	0	0	0	0	0	0	0	-
4	9	.02	0	0	0	0	0	0	1	0	0	0	-
5	9	.02	1	0	0	0	0	0	0	0	0	0	-
6	9	.02	0	0	0	0	1	1	1	1	1	1	5
7	9	.02	0	0	0	0	0	0	0	0	0	0	-
8	9	.02	0	0	0	0	0	0	0	0	0	0	-
9	9	.02	0	0	0	0	0	0	0	1	0	0	-
10	9	.02	1	0	0	0	0	0	1	1	0	1	7
11	9	.02	0	0	0	0	0	0	0	0	0	0	-
12	9	.02	1	1	1	0	1	0	1	0	1	1	1
13	9	.02	0	0	0	0	0	0	1	0	0	0	-
14	9	.02	0	0	0	0	0	0	0	0	0	0	-
15	9	.02	0	0	0	0	0	0	1	1	0	0	-
16	9	.02	0	0	0	0	0	1	1	1	1	1	6
17	9	.02	0	0	1	1	1	1	1	1	1	1	3
18	9	.02	0	0	0	0	0	1	0	0	0	0	-
19	9	.02	0	0	0	0	0	0	0	0	0	0	-
20	9	.02	0	0	1	0	0	0	0	0	0	0	-
1	9	.05	0	0	0	0	0	0	0	0	0	1	9
2	9	.05	0	1	0	0	0	1	1	1	1	1	6
3	9	.05	0	1	0	0	0	0	0	1	1	1	8
4	9	.05	0	1	0	0	0	1	1	0	1	1	2
5	9	.05	1	0	0	0	0	0	1	1	0	1	7
6	9	.05	0	0	0	0	0	0	1	1	1	1	7
7	9	.05	0	0	0	0	1	0	0	0	0	0	-
8	9	.05	0	1	0	0	0	0	0	0	1	0	-
9	9	.05	0	0	0	0	0	0	0	0	0	0	-
10	9	.05	0	0	0	0	1	1	1	0	1	1	5
11	9	.05	0	1	0	1	1	1	1	1	0	1	4
12	9	.05	0	0	0	0	0	0	0	0	0	0	-
13	9	.05	0	1	0	0	0	0	0	0	0	0	-
14	9	.05	0	0	0	0	0	1	0	0	0	0	-
15	9	.05	0	1	1	1	1	1	1	1	1	1	2
16	9	.05	0	0	0	0	0	0	0	0	0	0	-
17	9	.05	0	0	0	0	0	0	0	0	0	0	-
18	9	.05	0	1	1	0	0	0	0	0	0	0	-
19	9	.05	0	0	0	1	0	1	0	1	0	0	-
20	9	.05	0	0	0	0	0	0	0	0	0	1	10

1	1.5	.10	0 0 0 0 1 1 1 1 1 1	-	1	3	.10	1 0 0 0 0 0 1 0 0 0	-	1	9	.10	0 0 0 0 0 0 0 1 1 1	8
2	1.5	.10	0 1 0 0 0 0 1 0 1 1	4	2	3	.10	0 0 1 0 1 1 0 1 0 0	-	2	9	.10	0 0 0 0 0 0 0 0 0 0	-
3	1.5	.10	0 0 1 0 1 0 0 1 0 1	-	3	3	.10	0 1 0 0 0 0 1 0 0 0	-	3	9	.10	0 1 1 1 0 1 0 1 0 1	2
4	1.5	.10	0 1 1 0 0 1 1 1 0 1	4	4	3	.10	1 1 1 1 1 1 1 1 1 1	2	4	9	.10	0 0 0 1 1 1 0 1 1 1	4
5	1.5	.10	1 0 0 1 0 0 1 0 0 0	-	5	3	.10	1 0 0 1 1 1 1 1 1 1	3	5	9	.10	0 1 1 1 1 1 0 1 1 0	2
6	1.5	.10	1 0 0 0 0 1 1 0 1 0	7	6	3	.10	0 1 1 0 1 1 1 0 1 1	2	6	9	.10	1 1 1 1 1 1 1 1 1 0	1
7	1.5	.10	1 0 0 1 1 0 1 1 0 1	-	7	3	.10	1 1 1 1 1 0 1 0 0 1	1	7	9	.10	0 0 0 0 0 1 0 0 0 0	-
8	1.5	.10	1 1 0 1 1 1 1 0 1 0	1	8	3	.10	1 1 0 0 0 1 0 1 1 0	-	8	9	.10	0 0 0 0 0 0 0 0 0 0	-
9	1.5	.10	0 1 0 1 0 0 1 0 0 1	-	9	3	.10	0 0 1 1 1 1 1 0 1 1	2	9	9	.10	0 0 0 0 1 1 1 1 1 1	3
10	1.5	.10	0 1 0 1 0 0 1 1 1 0	-	10	3	.10	1 1 0 0 0 0 0 0 0 0	-	10	9	.10	0 1 0 0 0 1 1 1 1 1	7
11	1.5	.10	0 1 0 1 0 1 1 1 0 0	4	11	3	.10	0 0 1 1 1 1 1 1 1 1	3	11	9	.10	1 1 1 1 1 1 1 1 1 1	2
12	1.5	.10	1 1 1 0 0 0 0 1 0 1	8	12	3	.10	0 1 0 1 0 1 1 1 1 1	1	12	9	.10	0 0 0 0 0 0 0 0 1 1	9
13	1.5	.10	0 0 0 0 0 1 0 0 0 1	10	13	3	.10	1 0 0 1 1 0 1 1 1 0	-	13	9	.10	0 0 0 0 0 1 0 0 0 0	-
14	1.5	.10	0 1 1 1 1 1 1 1 1 0	2	14	3	.10	0 0 0 1 1 1 1 1 1 1	3	14	9	.10	0 0 0 1 1 1 1 1 1 0	4
15	1.5	.10	1 0 1 0 0 1 1 0 1 0	-	15	3	.10	0 0 0 1 0 0 0 0 0 0	-	15	9	.10	0 1 1 1 1 1 1 1 1 1	1
16	1.5	.10	1 1 0 0 1 0 0 0 1 0	7	16	3	.10	0 1 1 1 1 0 1 1 1 1	3	16	9	.10	0 1 1 1 1 1 1 1 1 1	2
17	1.5	.10	1 0 0 0 1 1 0 1 0 1	8	17	3	.10	0 0 1 1 1 1 1 0 1 0	4	17	9	.10	0 1 0 0 0 0 0 0 0 1	-
18	1.5	.10	0 1 1 0 0 1 0 0 1 0	6	18	3	.10	0 0 0 1 1 1 1 1 1 0	5	18	9	.10	0 0 0 0 0 0 0 0 0 0	-
19	1.5	.10	1 0 1 0 1 0 1 1 0 0	2	19	3	.10	0 0 0 0 1 1 1 1 1 1	6	19	9	.10	0 0 1 0 0 0 0 0 1 0	-
20	1.5	.10	0 0 0 1 1 0 1 1 1 0	8	20	3	.10	0 0 0 0 0 0 0 0 1 1	9	20	9	.10	0 0 0 0 0 0 0 0 0 0	-
1	1.5	.20	1 0 1 1 0 0 0 1 0 1	1	1	3	.20	0 0 1 0 0 0 0 0 1 0	-	1	9	.20	0 0 0 1 0 0 0 0 0 0	-
2	1.5	.20	0 0 1 0 0 0 1 1 1 0	3	2	3	.20	1 0 0 1 0 1 1 1 1 0	3	2	9	.20	0 0 1 0 0 1 0 1 1 1	6
3	1.5	.20	1 0 0 1 0 1 0 0 1 1	4	3	3	.20	1 0 1 1 1 1 0 0 0 1	1	3	9	.20	0 0 0 1 0 0 1 0 0 0	-
4	1.5	.20	1 1 0 1 0 1 1 1 1 0	3	4	3	.20	1 0 0 0 0 1 1 1 1 0	4	4	9	.20	0 0 1 1 1 1 1 1 1 1	4
5	1.5	.20	1 0 1 0 0 1 0 0 0 0	9	5	3	.20	1 1 0 0 1 1 0 1 1 1	5	5	9	.20	0 1 1 1 1 1 1 0 0 1	2
6	1.5	.20	0 0 1 0 1 0 1 0 1 1	2	6	3	.20	0 0 1 0 0 1 0 1 0 1	7	6	9	.20	0 1 1 1 1 0 0 1 1 1	2
7	1.5	.20	0 1 1 1 1 1 1 0 1 0	5	7	3	.20	1 1 1 1 0 1 1 1 0 1	1	7	9	.20	1 1 1 1 1 1 1 1 0 1	1
8	1.5	.20	0 0 1 1 1 0 1 1 0 1	2	8	3	.20	0 1 0 1 1 1 1 1 1 1	2	8	9	.20	0 0 0 1 1 1 1 1 1 1	2
9	1.5	.20	0 0 1 1 0 0 1 1 0 1	2	9	3	.20	0 1 1 0 0 1 1 1 0 0	1	9	9	.20	0 0 1 0 0 0 0 0 0 0	-
10	1.5	.20	0 0 0 1 0 1 1 1 0 1	4	10	3	.20	1 1 1 1 1 1 1 1 1 0	2	10	9	.20	0 0 0 0 0 0 1 1 1 1	7
11	1.5	.20	1 1 0 1 1 0 1 1 0 0	4	11	3	.20	0 0 1 1 0 0 1 0 1 1	1	11	9	.20	0 1 1 0 1 1 1 1 1 1	2
12	1.5	.20	0 1 1 1 0 0 1 0 1 1	3	12	3	.20	0 1 1 0 1 0 1 0 0 1	-	12	9	.20	0 0 0 1 1 1 1 1 1 1	3
13	1.5	.20	0 1 1 1 1 1 0 1 0 1	1	13	3	.20	1 1 1 1 1 1 0 1 1 0	1	13	9	.20	0 1 1 1 1 1 0 0 1 1	2
14	1.5	.20	0 0 1 0 0 1 1 1 0 1	10	14	3	.20	0 0 0 0 1 1 1 0 1 0	5	14	9	.20	0 0 1 1 1 1 1 1 1 1	3
15	1.5	.20	0 0 0 1 0 1 0 1 0 1	1	15	3	.20	0 1 0 1 1 1 1 1 1 1	4	15	9	.20	0 0 0 0 0 0 0 1 1 0	8
16	1.5	.20	0 1 0 0 0 0 0 1 0 1	1	16	3	.20	1 0 0 0 0 0 0 0 1 1	9	16	9	.20	1 1 1 0 1 1 1 1 1 1	1
17	1.5	.20	0 1 1 0 1 0 0 0 1 0	1	17	3	.20	1 0 1 0 1 1 0 0 1 1	3	17	9	.20	0 0 0 1 0 0 0 1 1 1	8
18	1.5	.20	1 0 0 1 0 1 0 1 1 1	9	18	3	.20	0 0 1 1 0 0 1 1 1 1	2	18	9	.20	1 1 1 1 1 1 1 1 1 1	1
19	1.5	.20	0 1 0 1 1 1 0 1 1 0	-	19	3	.20	0 1 1 1 0 1 1 1 1 1	3	19	9	.20	0 1 1 1 1 1 1 1 1 1	4
20	1.5	.20	0 0 0 0 1 1 0 0 1 0	-	20	3	.20	0 0 0 0 1 1 1 1 1 1	5	20	9	.20	0 0 0 0 0 0 1 1 1 1	8

Table A1. Stimuli from the learning experiment (0 indicates blue ball, 1 indicates red ball) and period in which regime shifted to blue, if ever.