Habits as Adaptations: An Experimental Study*

Ludmila Matysková[†]
University of Bonn

Brian Rogers[‡]
Washington University in St. Louis

Jakub Steiner[§]
University of Zurich and CERGE-EI

Keh-Kuan Sun[¶]
Washington University in St. Louis

September 2, 2019

Abstract

When observable cues correlate with optimal choices, habit-driven behavior can alleviate cognition costs. We experimentally study the degree of sophistication in habit formation and cue selection. To this end, we compare lab treatments that differ in the information provided to subjects, holding fixed the serial correlation of optimal actions. We find that a particular cue–own past action–affects behavior only in treatments in which this habit is useful. The result suggests that caution is warranted when modeling habits via a fixed non-separable utility. Despite their sophistication, lab behavior also reveals myopia in information acquisition.

keywords: habit formation, rational inattention.

JEL codes: C91, D8, D9

*This paper builds on a dissertation chapter of Ludmila Matyskova entitled "Habit Formation: An Experimental Study". We thank the audience in Cognition and Decision Lab Meeting at Columbia University, participants at the Midwest Economics Association Meeting 2019, the Economics Graduate Student Conference 2018, 2018 European Winter Meeting of the Econometric Society, 2019 Sloan-Nomis Workshop, and EEA-ESEM 2019. We thank Andrew Caplin, Mark Dean, Emir Kamenica, Colin Stewart, and Michael Woodford for comments. Matyskova acknowledges the funding CRC TR 224, Project B02, provided by the German Research Foundation (DFG). Steiner acknowledges the financial support of the Czech Science Foundation grant 16-00703S and the ERC grant 770652.

[†]email: ludmila.matyskova@cerge-ei.cz
[‡]email: brogers@wustl.edu
[§]email: jakub.steiner@cerge-ei.cz
[¶]email: sun.k@wustl.edu

1 Introduction

Habits play an important role in economic discourse. Economists employ them to explain diverse phenomena ranging from inertia of consumption to brand loyalty. In many cases, the standard modeling approach represents habits via a fixed time-nonseparable utility function, thus leaving the issues of when and why habits form, and their responses to counterfactual environments, unaddressed.[1] Psychologists offer a view on both the purpose of and the mechanism underlying habit formation. They typically define habits as automated responses triggered by cues, where cues are elements of the decision history that empirically correlate with optimal continuation choices. In this view the purpose of habits is to alleviate cognition costs; see Andrews (1903), Lally et al. (2010), and Wood & Neal (2007).

We ask whether habits originate by automatically following a cue that empirically correlates with well-performing choices or whether, instead, habits are second-best adaptations. Our main experimental result mostly supports the latter hypothesis. We view the finding as good news for the predictability of habits. An understanding of habit formation rooted in optimization can inform analysts which cues, out of several available cues, a decision-maker will leverage. Modeling habits as optimal adaptations also permits counterfactual predictions of habit strength under various policies.

To discriminate between the automatic and optimization origins of habits, we compare treatments from a lab experiment in which subjects face a sequence of tasks generated by a given stochastic process. The compared treatments differ only in the information feedback. If habits form automatically whenever past cues and optimal continuation choices correlate, then the variation in feedback should not impact habit formation and the selection of cues. Our data, however, show that subjects form distinct habits across these treatments; moreover, the cues selected are naturally rationalized as adaptations to the information provided.

In our experiments, the basic task confronted by subjects is to recognize a binary state variable presented visually on a computer screen. Correctly identifying the state requires moderate cognitive effort.[2] Subjects face two periods of this state-recognition task, across which the state evolves according to a known stochastic process with positive serial correlation. In the treatment without feedback, we reveal both realized states to the subjects only at the conclusion of the two-period sequence. We find that subjects form a habit in this

[1] Time-nonseparable utility provides a good representation of preferences in the presence of real switching costs, addiction, etc. In this paper, and through our experimental design, we abstract from such direct causes of action inertia.

[2] Identifying the state amounts to conducting a counting process, so that we can plausibly assume that frictions in the cognitive process are the main source of errors.

treatment: the first-period outcome predicts the second-period choice (controlling for the second-period state) in this treatment. The cue that subjects leverage is their first-period *action*; the first-period state does not predict the second-period action. In other words, the behavioral pattern exhibits action inertia. The habit alleviates the subject's cognitive burden since, due to the serial correlation of the states, the first-period action contains useful information about the second-period optimal choice, and the subjects utilize this information.

In the other treatment, with information feedback, we employ the same state-generating process, but we reveal the first-period state in between the two periods. Subjects again form a habit in this treatment: payoff-irrelevant elements of the history predict the continuation choice (controlling for the second-period state). However, importantly, the cue changes relative to the previous treatment. The first-period action is no longer predictive; all of the predictive power is associated with the first-period state, which contains superior information about the optimal continuation action relative to the first-period action. This finding is inconsistent with the view that habits arise as an automatic consequence of the correlation between cues and optimal actions. Rather, the result suggests that our subjects choose cues optimally, according to their informational content. As a further check, we ran additional treatments (with and without information feedback) in which the states were serially independent. As expected, subjects do not form habits in these treatments; the second-period choice is independent of all first-period variables.

If habits are, as our results suggest, optimal adaptations, then their strength should vary predictably with the decision environment, in particular, with the incentive stakes and the serial correlation of states. When stakes are decreased or correlation increased, the trade-off between reliance on the cues and the acquisition of new information shifts in favor of the cues. Thus, we predict that habits become stronger—cues become more predictive of continuation behavior—when stakes are lower and correlation is greater. We test this hypothesis experimentally. For the correlated treatments, changes in stakes and correlation have no impact on the cue selection, but they do affect the strength of habits. We obtain strong statistical evidence in favor of the predicted comparative statics when the selected cue is the past action. When the cue is the past state, the evidence continues to support the prediction, although it is less conclusive.

Although our subjects use information, once they acquire it, sophisticatedly, we also find indirect evidence of myopia in the information-acquisition choices. When states are correlated and feedback is not provided then information is relatively more valuable in the first period since it is useful in both periods. Therefore, a forward-looking decision-maker should acquire more information in the first than in the second period, and this should be

manifested in the observed accuracy of choices. Since we do not observe differences in the accuracy of choice across periods, we conjecture that the subjects do not fully internalize the continuation value of information.

We supplement the experiments with a model that derives habit formation from primitive assumptions on the information-processing friction. In the model, a decision-maker chooses information structures (i.e., a strategy for how to acquire information about the states) and trades off the accuracy of her information against an acquisition cost. The model allows us to formalize the above intuitive predictions about habit formation, cue selection, and the comparative statics of habit strength.

Popular macroeconomic models explain the empirically observed inertia of consumption by imposing a time-nonseparable utility function $u\left(c_t - c^{t-1}\right)$, where c^{t-1} is an aggregate of the consumption history, e.g. Pollak (1970) and Abel (1990). When u is concave, high past consumption triggers high current consumption; i.e., c^{t-1} becomes the cue for a consumption habit. Since the assumed utility representation is exogenous, the modeling choice of the aggregate c^{t-1} is not obvious and specifications in the literature include aggregates of past population-wide consumption, past individual consumption, and past individual consumption of specific categories of goods; see Schmitt-Grohé & Uribe (2007) for a review. While time-nonseparable utility functions capture the causes of action inertia in applications with adaptation frictions or with dependencies akin to smoking, this paper focuses on an alternative explanation that is compatible with time-separable payoffs.

Laibson (2001) proposes a model of habit formation rooted in psychological forces that, like ours, focuses on an endogenous selection among several available habit cues, albeit, unlike in our case, the cue selection is not rooted in the optimization of cognition costs. Camerer et al. (n.d.) study a model of habit formation inspired by neuroeconomics and advocate for an optimization-based origin of habits. Angeletos & Huo (2018) prove the observational equivalence between a model featuring higher-order uncertainty, and a model of a representative agent with consumption habits.

Our model of habit formation belongs to the rational-inattention literature originating in Sims (2003). It builds, in particular, on the discrete dynamic rational-inattention model by Steiner et al. (2017), which in turn extends a static model by Matějka & McKay (2015). Rational inattention models have been used to derive inertia of behavior in a macroeconomic context, see Mackowiak & Wiederholt (2009) for a theoretical contribution and Khaw et al. (2017) for an experimental exploration. Khaw & Zorrilla (2018) find that lab subjects develop action inertia in a consumption/saving problem and, like us, the authors interpret the observed inertia as habits.

2 Habits and cues

2.1 Definitions

We study habit formation in the simplest possible setting. A decision-maker (DM) chooses a binary action $a_t \in \{0, 1\}$ in each of two periods to maximize $\sum_{t=1}^{2} u(a_t, \theta_t)$. The binary state $\theta_t \in \{0, 1\}$ evolves according to a stochastic process known to the DM. The first-period state attains value 1 with prior probability p_1, and $\theta_2 = \theta_1$ with probability $\gamma \geq \frac{1}{2}$ for each value of θ_1. The two states are independent when $\gamma = \frac{1}{2}$, and they are positively correlated if $\gamma > \frac{1}{2}$. In the latter case, we say that the states are persistent. The DM's task in each period is to match the action to the state; $u(a, \theta) = s$ if $a = \theta$ and zero otherwise; $s > 0$ is the stake.

An analyst collects data on the states and actions across many repetitions of the two-period sequence. In its idealized form, the analyst observes the joint probability distribution $\pi(\theta_1, a_1, \theta_2, a_2)$ over the quadruples of states and chosen actions. Our data extend the state-dependent stochastic-choice data introduced by Caplin & Dean (2015) in a static setting to the dynamic context considered here.

We say that the DM *forms a habit* if there exists a triple $(\theta_1, \theta_2, a_1) \in \{0, 1\}^3$ such that $\pi(a_2 \mid \theta_2, \theta_1, a_1) \neq \pi(a_2 \mid \theta_2)$. Otherwise, if a_2 is independent of (θ_1, a_1) conditional on θ_2, we say that the DM *does not form a habit*. Thus, the DM forms a habit if the history of the process – which is irrelevant to the continuation payoff – predicts continuation behavior. Our definition of habits is behavioral in nature and distinct from the commonly used non-separable utility approach. Our analyst knows that the DM's utility is, in fact, time-separable; she attributes any correlation between the history and continuation behavior to imperfections in the decision process, and refers to the predictive power of the history as a habit.

Additionally, we define cues that drive the habitual behavior. Is the habitual behavior in the second period, if it arises, triggered by the past state θ_1, or by the past action a_1? Let z be one of the two random variables in the set $\{\theta_1, a_1\}$ and w be the complementary variable from this set. We say that z *is the cue for the habit* if (i) $\pi(a_2 = 1 \mid \theta_2, z = 1) > \pi(a_2 = 1 \mid \theta_2, z = 0)$, and if (ii) $\pi(a_2 = 1 \mid \theta_2, z, w) = \pi(a_2 = 1 \mid \theta_2, z)$. Thus, for instance, the past action a_1 is the cue for the habit if the probability that the DM chooses the high action in the second period increases with a_1 given θ_2, and θ_1 has no additional predictive power. The latter condition prevents a spurious identification of cues. Since θ_1 and a_1 may be correlated (and indeed are correlated in our data), it may happen that they both correlate with continuation behavior, but all the predictive information is contained in only one of

them.[3]

Habits exhibit a continuous range of strength since the correlation between cues and continuation behavior varies with the DM's environment. We capture this as follows. Suppose that the DM has developed a habit with cue $z \in \{\theta_1, a_1\}$. For a state value $\theta \in \{0, 1\}$, we define the *habit strength* $\phi_z(\theta)$ at state realization $\theta_2 = \theta$ to be

$$\phi_z(\theta) = \frac{\pi(a_2 = \theta \mid \theta_2 = \theta, z = \theta)}{\pi(a_2 = \theta \mid \theta_2 = \theta, z = 1 - \theta)},$$

which measures how strongly the probability that the DM chooses the correct a_2 varies with the cue value in the state $\theta_2 = \theta$.

2.2 Hypotheses

Our first set of hypotheses predict that habits are formed only if available cues contain useful information about optimal continuation choice and that, in such cases, the most informative available cue is selected.

Hypothesis.

(H1.1) If states are independent, then the DM does not form a habit.

(H1.2) If states are persistent and θ_1 is not revealed then a habit is formed with the cue a_1.

(H1.3) If states are persistent and θ_1 is revealed then a habit is formed with the cue θ_1.

The next hypothesis is based on the observation that when states are correlated and the first-period state is not revealed, information acquired in the first period has a continuation value. Thus, forward-looking subjects should acquire more information in the first period relative to the second one, as the early information is useful twice.

Hypothesis (*H2*)**.** *If the states are persistent and θ_1 is not revealed, then the first-period response is correct with a higher probability than the second-period response.*

Finally, to test the comparative-statics predictions arising from the DM's optimization problem, we study two specifications of stakes and state correlations. Stakes are low and state persistence is high in the *strong-habit* treatments, whereas stakes are high and persistence is low in the *weak-habit* treatments. We expect the natural comparative statics across the weak and strong pairs of treatments.

[3]Our definition of cues has roots in cognitive psychology and neuroscience. These fields conceptualize habits as tendencies to choose actions that have been previously rewarding (e.g. Dezfouli & Balleine 2012). Since our first-period cues correlate with the optimal first-period actions, conditioning on the cues correlates the second-period choices with the first-period optimal choices.

Weak treatments			Strong treatments		
high stakes	*no feedback*	*feedback*	*low stakes*	*no feedback*	*feedback*
state independence	no habit	no habit	*state independence*	no habit	no habit
low persistence	weak habit with cue a_1	weak habit with cue θ_1	*high persistence*	strong habit with cue a_1	strong habit with cue θ_1

Table 1: Hypotheses on habit formation for the eight experimental treatments.

Hypothesis (*H3*)**.** *For each treatment with persistent states, the formed habits are stronger when persistence is high and stakes are low than when persistence is low and stakes are high.*

In summary, the experimental treatments vary along three dimensions: (i) we consider independent or positively serially correlated states, (ii) we reveal or do not reveal θ_1 before the second period, and (iii) we vary the stakes and the state correlation. Table 1 summarizes our hypotheses on habit formation, cue selection, and habit strength in the resulting eight treatments.

3 A rational-inattention model of habit formation

The model we present here explains habits as optimal second-best adaptations to an information-processing friction. The DM solves the two-period binary decision problem with an evolving state and time-separable utility from the beginning of Subsection 2.1. The DM conducts costly statistical experiments that produce signals x_t in periods $t = 1, 2$. Additionally, in between periods 1 and 2, she receives an exogenous signal y. In each period t, she chooses an action according to a pure action strategy σ_t that maps the observed signals up to period t to a_t; $a_1 = \sigma_1(x_1)$ and $a_2 = \sigma_2(x_1, x_2, y)$. The DM controls the experiments that generate x_t and can condition the employed experiment on all the available information at the given period: Let X, $|X| \geq 2$, be a fixed signal space. The DM can choose any experiment $f_1(x_1 \mid \theta_1)$ and any system of experiments $f_2(x_2 \mid \theta_2, x_1, y)$ that govern the conditional probability distribution of the signals $x_t \in X$ for each combination of the values of the random variables specified in the condition. The signal x_1 is constrained to be independent from θ_2 conditional on θ_1. Similarly, x_2 is constrained to be independent from θ_1 conditional on (θ_2, x_1, y); the DM learns about θ_t only in period t.

We consider two distinct processes that generate the exogenous signal y. In one case, y perfectly reveals the first state; $y = \theta_1$, and we say that the *DM receives feedback*. In the other case, $y = y_0$, where y_0 is an arbitrary constant, and we say that the *DM does not receive feedback*.

Following the rational-inattention literature, we specify a uniformly posterior-separable information cost; see Caplin et al. (2017). Let $H : [0,1] \longrightarrow \mathbb{R}$ be a strictly concave function. The cost of the first-period experiment $f_1(x \mid \theta)$ is

$$I(\theta_1; x_1) = \mathrm{E}_{\hat{q}_1}\left[H(p_1) - H\left(\hat{q}_1\right)\right]$$

where p_1 is the prior probability that $\theta_1 = 1$ and $\hat{q}_1 = \Pr\left(\theta_1 = 1 \mid x_1\right)$. Together with the literature, we interpret $H(q)$ as the measure of uncertainty of the belief that assigns probability q to $\theta_1 = 1$ and $I(\theta_1; x_1)$ is the expected reduction of the uncertainty achieved by the observation of the signal x_1. When H equals Shannon entropy, then $I(\theta_1; x_1)$ is the mutual information between the random variables θ_1 and x_1, and the information cost coincides with that in Sims (2003). Analogously, for any given signal realizations x_1 and y, the cost of the second-period experiment $f_2(x_2 \mid \theta_2, x_1, y)$ is

$$I(\theta_2; x_2 \mid x_1, y) = \mathrm{E}_{\hat{q}_2}\left[H(p_2) - H\left(\hat{q}_2\right) \mid x_1, y\right],$$

where $p_2 = \Pr\left(\theta_2 = 1 \mid x_1, y\right)$ is the DM's belief about θ_2 from the beginning of the second period and $\hat{q}_2 = \Pr(\theta_2 = 1 \mid x_1, y, x_2)$ is the posterior belief. The DM maximizes her expected net payoff:

$$\max_{f_1, f_2, \sigma_1, \sigma_2} \mathrm{E}\left[u\left(\sigma_1\left(x_1\right), \theta_1\right) + u\left(\sigma_2\left(x_1, x_2, y\right), \theta_2\right) - I\left(\theta_1; x_1\right) - I\left(\theta_2; x_2 \mid x_1, y\right)\right].$$

Let $\pi(\theta_1, a_1, \theta_2, a_2)$ be the joint distribution of the states and actions generated by the optimal experiments f_1^* and f_2^* and action strategies σ_1^*, σ_2^*. We impose a regularity condition that all quadruples $(\theta_1, a_1, \theta_2, a_2)$ are attained with positive probabilities. The condition holds when the cost function is sufficiently scaled down, and is satisfied in our experimental data.

Lemma 1. *The optimal joint distribution π of states and actions is unique.*

Proofs can be found in the Appendix.

The next proposition formalizes Hypothesis 1 in the context of our model.

Proposition 1.

(P1.1) If states are independent, then the DM does not form a habit.

(P1.2) If states are persistent and the DM does not receive feedback, then she forms a habit with the cue a_1.

(P1.3) If states are persistent and the DM receives feedback, then she forms a habit with the cue θ_1.

We expect P1.1 and P1.3 to hold for a wide range of information-processing frictions. In contrast, part P1.2 relies in a subtle way on the assumption of a posterior-separable cost. It may fail for alternative information-acquisition specifications, such as those in which the DM pays for a reduction of Gaussian noise. Recall from our definition that a_1 is the habit cue if, conditionally on (θ_2, a_1), the state θ_1 does not predict a_2. This is indeed the case under the posterior-separable cost function because each action $a_1 = 0, 1$ corresponds to a unique information set; without feedback, a_1 is a sufficient statistics for the DM's information from the beginning of the second period. In alternative models, such as the Gaussian one, each $a_1 = 0, 1$ may correspond to multiple information sets. Then, θ_1 may predict the DM's information set from the beginning of the second period, and hence indirectly predict a_2, controlling for (θ_2, a_1).

We turn now to the comparative-statics results, which are counterparts of Hypotheses 2 and 3. To proceed, we impose additional assumptions on the information-cost specification.

Assumption A. *The information-cost function satisfies the following four properties:*
(A.1) symmetry; $H(q) = H(1-q)$ for all $q \in [0, 1]$,
(A.2) it is twice differentiable,
(A.3) it is steep at certainty; $\lim_{q \to 0^+} H'(q) = \infty$,
(A.4) $H''(q) \leq H''(q')$ for any q, q' such that $|q - 1/2| > |q' - 1/2|$.

In particular, all four properties in Assumption A are satisfied by the Shannon entropy cost and by the log-likelihood cost from Morris & Strack (2017).

We now present the counterpart of Hypothesis 2.

Proposition 2 (*P2*)**.** *Suppose that the cost function satisfies the properties in Assumption A. If the states are persistent and the DM does not receive feedback, then the DM chooses the correct action with higher probability in period 1 than in period 2,*

$$\pi(a_1 = \theta_1) > \pi(a_2 = \theta_2).$$

The following comparative-statics results are counterparts of Hypothesis 3. The proof of one of them–comparative statics with respect to stakes in the setting without feedback, (iii)–exploits the functional form of the Shannon-entropy cost. The other comparative statics results are derived for all cost functions satisfying Assumption A. Recall that Shannon entropy for a binary random variable is given by $H_{\text{Shannon}}(q) = -q \log q - (1-q) \log(1-q)$.

Proposition 3 (*P3*)**.** *Suppose the states are persistent and the cost function satisfies Assumption A. Then, the habit strength*

(i) increases with the state persistence both with and without feedback,

(ii) decreases with stakes with feedback,

(iii) if the cost function equals the Shannon entropy, then the habit strength decreases with stakes without feedback too.

4 Experimental design and data

Our experimental design follows Caplin & Dean (2015). Subjects were presented with images of a 10×10 matrix of red and blue dots on a computer screen. In each matrix, either 51 red and 49 blue (state $\theta = 1$) or 51 blue and 49 red dots ($\theta = 0$) are displayed. The positions of the colored dots are random conditional on the state; see the screenshot in Appendix A.3. Subjects are incentivized to determine the majority color and do not face any explicit information cost; any perception errors stem from frictions in the cognitive process. When a subject is ready, she enters her choice by clicking one of two radio buttons marked "Red" and "Blue".[4] To ensure a reasonable duration of the experiment, each image disappeared after 45 seconds.[5] The experiments were implemented using z-tree (Fischbacher 2007). We refer to the above one-period decision problem as the *counting task*.

We recruited 76 subjects from the University of California, Santa Barbara over 4 sessions during May 2018. In each session, subjects faced 4 treatments. Each treatment consisted of 12 iterations and each iteration consisted of the two-period decision problem described above, with one counting task per period. Thus, each subject faced $96 = 4 \times 12 \times 2$ counting tasks in total. At the conclusion of the session, the software randomly chose a single counting task for each subject, and the subject's payment was based only on the outcome of that task.

An "iteration" is our basic unit of observation. In each iteration, both state realizations were equally likely in the first period: $p_1 = 1/2$. The four treatments per session are defined by the combinations of (i) the state persistence, where I denotes *independent* and C denotes *correlated* states, and (ii) whether θ_1 was revealed in between the two periods, where F denotes the provision of information feedback and N denotes no provision. In addition, in two of the sessions we used parameters that we hypothesized to induce strong habits; treatments

[4]Since we are interested in serial correlations that arise in the absence of real switching costs, we set the positions of the blue and red radio buttons to randomly vary across tasks. Thus, provision of the same answer in consecutive periods does not arise from a mental or physiological advantage.

[5]The subjects could submit their choices after the image had disappeared. The time constraint was binding at similar levels across treatments, in only 2-3% of problems.

Weak treatments			Strong treatments		
high stakes: \$10	*no feedback*	*feedback*	*low stakes: \$7*	*no feedback*	*feedback*
persistence $\gamma = .5$	INW	IFW	*persistence* $\gamma = .5$	INS	IFS
persistence $\gamma = .75$	CNW	CFW	*persistence* $\gamma = .9$	CNS	CFS

Table 2: Experimental treatments.

in these two sessions are denoted by S, and the treatments in the other two sessions are labeled by W. We thus have 8 treatments from the set $\{I, C\} \times \{N, F\} \times \{W, S\}$; see Table 2. We have randomly generated the state sequence once for each of the 8 treatments and used it in both sessions in which the treatment occurred. Within a treatment, each subject faced the same sequence of images.[6] Additionally, we ran a preliminary session prior to the 4 regular sessions. The results from this session are consistent with those from the regular sessions. We omit this session from the analysis due to a minor error in the experimental procedure; see Appendix A.2.

Each session lasted approximately 90 minutes. In all cases, subjects received a \$10 show-up fee. Average total earnings per subject were \$17.27 paid in cash at the conclusion of the experiment. The expected additional payoff for each correct answer was $\$\frac{10}{96} \approx \$.1$ in W treatments and it was $\$\frac{7}{96} \approx \$.07$ in S treatments which is comparable to incentives in Caplin & Dean (2013) that varied from \$.01 to \$.15. See Appendix A.3 for our experimental instructions.

5 Results

We present basic summary statistics in Table 3. The aggregate accuracy of choices is high and homogeneous across treatments and periods. Accuracy is heterogenous at the individual level; the number of correctly answered tasks per subject varied from 46 to 96 out of 96 tasks. (Mild action persistence in the treatment IFW, in which the frequency of $a_1 = a_2$ is 0.60, is caused by the realized frequency of $\theta_1 = \theta_2$ being 0.67 and by the subjects' attentiveness to the state realizations.[7])

We proceed to test for the presence of habits and to identify the cues. To examine how θ_1 and a_1 predict a_2, we run separate logit regressions for all 8 treatments of the form:

[6] The treatments were ordered either IF, CF, IN, CN or IN, CN, IF, CF, once each for the two strong (S) and two weak (W) parameter sessions.

[7] The frequency .67 of $\theta_1 = \theta_2$ corresponds to 8 cases out of 12, which occurs with an a priori probability $\binom{12}{8}.5^8.5^4 \approx .12$.

	Frequency of $a_1 = \theta_1$	Frequency of $a_2 = \theta_2$	Frequency of $a_2 = a_1$	Frequency of $a_1 = \theta_2$
INW	.84	.86	.51	.50
IFW	.85	.85	.60	.61
CNW	.87	.86	.78	.74
CFW	.89	.90	.78	.77
INS	.87	.85	.51	.50
IFS	.82	.82	.53	.55
CNS	.84	.85	.91	.84
CFS	.86	.87	.75	.73

Table 3: Data summary

$$a^n_{2,i} = \begin{cases} 1 & \text{if } \beta_0 + \beta_{\theta_2}\theta^n_2 + \beta_{\theta_1}\theta^n_1 + \beta_{a_1}a^n_{1,i} + \beta_{se}session + \beta_{sc}score^n_i + \beta_{sc\theta_2}score^n_i\theta^n_2 + \varepsilon^n_i > 0, \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

with robust standard errors clustered at the subject level, where $a^n_{t,i}$ is the action taken by subject i in iteration $n = 1, \ldots, 12$ and period $t = 1, 2$; θ^n_t is the realized state in iteration n and period t; and *session* is a dummy variable indicating session (each of the 8 treatments occurs in exactly two sessions).[8] Finally, we include the interaction term $score^n_i\theta^n_2$ (and the term $score^n_i$), where $score^n_i$ is a subject-specific proxy for counting ability. It is the total number of correct answers by subject i in all treatments (excluding the two choices from iteration n of the considered treatment to avoid endogeneity). The interaction term $score^n_i\theta^n_2$ captures the idiosyncratic sensitivity of the subjects to the variation in θ_2. Otherwise, failing to account for heterogeneity in counting ability may lead to spurious significance of θ_1 in the correlated treatments without feedback: for a given a_1, $\theta_1 = a_1$ predicts high counting ability which, in turn, predicts $a_2 = \theta_2$. Hence, under state persistence, it predicts $a_2 = \theta_1$.

Table 4 reports the estimated average marginal effects, their standard errors and p-values of the explanatory variables of interest. Below, we list conclusions that we draw from these estimates.

Our first result is an elementary confirmation that subjects pay attention to the payoff states.

Result (*R0*)**.** *Subjects pay attention to θ_2: θ_2 predicts a_2 in all treatments.*

The next findings confirm Hypotheses H1.1, H1.2, and H1.3.

Result.

[8] In the treatment CNS, the state realizations satisfied $\theta^n_1 = \theta^n_2$ for all n, and thus we dropped θ_1.

	IFW	INW	CFW	CNW
a_1	-.021 (.036) .548	.034 (.041) .406	.017 (.032) .603	**.191** (**.051**) .000
θ_1	.071 (.043) 092	-.026 (.049) .596	**.258** (**.058**) .000	.002 (.036) .948
θ_2	**.681** (**.032**) .000	**.692** (**.054**) .000	**.611** (**.046**) .000	**.629** (**.067**) .000

	IFS	INS	CFS	CNS
a_1	-.031 (.037) .401	.037 (.045) .416	-.033 (.026) .204	**.511** (**.110**) .000
θ_1	.009 (.040) .817	-.034 (.044) .430	**.498** (**.098**) .000	omitted
θ_2	**.632** (**.045**) .000	**.700** (**.036**) .000	**.425** (**.121**) .000	**.367** (**.098**) .000

Table 4: Average marginal treatment effects, (their standard errors), and the p-values in the second lines. Bold values indicate significance at the 1% level.

(R1.1) When the states are independent, the subjects do not form habits: neither a_1 nor θ_1 predict a_2 in treatments IFW, INW, IFS, and INS.

(R1.2) When the states are persistent and feedback is not provided, the subjects form a habit with cue a_1: a_1 predicts and θ_1 does not predict a_2 in treatments CNW and CNS.

(R1.3) When the states are persistent and the feedback is provided, the subjects form a habit with cue θ_1: θ_1 predicts and a_1 does not predict a_2 in treatments CFW and CFS.

Next, we compare choice accuracy in the first and the second periods both in treatments CNW and CNS. For $t = 1, 2$, let $r_t \in \{c, w\}$ indicate correct and wrong response in period t, respectively. We test the hypothesis that the marginal probabilities of correct answers are equal across the two periods, $\Pr(r_1 = c) = \Pr(r_2 = c)$. Since r_t are binary and correlated random variables, we apply McNemar's test; see McNemar (1947). We cannot reject the hypothesis of equal precisions both for the CNW treatment (p-value .29) and for the CNS treatment (p-value .42). Accordingly, the data do not confirm Hypothesis H2. We interpret this finding as indirect evidence of myopia in information acquisition, whereby subjects do not fully account for the positive continuation value of information under state persistence.

Result (*R2*)**.** *Choice accuracy does not differ significantly across the first and second period of the treatments with persistent states without feedback.*

To analyze the comparative statics of habit strength, we focus on the four treatments with persistent states in which habits are formed, and we compare the habit strength across the weak and strong treatments. Namely, for the treatments without feedback, we pool the data from CNW and CNS and create a dummy variable $\delta \in \{0, 1\}$ to indicate treatment S. We run the same logit specification as in (1) with the inclusion of the additional variables δ, $\delta\theta_2^n$, $\delta a_{1,i}^n$, $\delta score_i^n$, and $\delta score_i^n \theta_2^n$.[9] Since empirically the habit cue is a_1, we estimate the difference between the average marginal effect of a_1 conditional on $\delta = 1$ (S) and its average marginal effect conditional on $\delta = 0$ (W),

$$\begin{aligned}\Delta_{CN} &= \mathrm{E}_X \left[\Pr(a_2 = 1 \mid a_1 = 1, X, \delta = 1) - \Pr(a_2 = 1 \mid a_1 = 0, X, \delta = 1)\right. \\ &- \left.\big(\Pr(a_2 = 1 \mid a_1 = 1, X, \delta = 0) - \Pr(a_2 = 1 \mid a_1 = 0, X, \delta = 0)\big)\right],\end{aligned}$$

where X stands for all explanatory variables other than a_1 and δ. We obtain a point estimate $\hat{\Delta}_{CN} = .31$ with standard error .12, which is highly significant (p-value .009).

Analogously, for the treatments with feedback, we pool the data from treatments CFW and CFS and create a dummy variable $\delta \in \{0, 1\}$ to indicate the strong treatment. We run the regression model (1) with the inclusion of δ, $\delta\theta_2^n$, $\delta\theta_1^n$, $\delta a_{1,i}^n$, $\delta score_i^n$, and $\delta score_i^n \theta_2^n$. Since the habit cue is θ_1, we estimate the difference between the average marginal effect of θ_1 conditional on $\delta = 1$ (S) and its average marginal effect conditional on $\delta = 0$ (W),

$$\begin{aligned}\Delta_{CF} &= \mathrm{E}_X \left[\Pr(a_2 = 1 \mid \theta_1 = 1, X, \delta = 1) - \Pr(a_2 = 1 \mid \theta_1 = 0, X, \delta = 1)\right. \\ &- \left.\big(\Pr(a_2 = 1 \mid \theta_1 = 1, X, \delta = 0) - \Pr(a_2 = 1 \mid \theta_1 = 0, X, \delta = 0)\big)\right].\end{aligned}$$

Here, we obtain the point estimate $\hat{\Delta}_{CF} = .23$ with standard error .12, which is marginally significant (p-value .06).

Result (*R3*)**.** *We find conclusive (inconclusive) statistical evidence that the habit formed in the correlated treatments without (with) feedback is stronger in the treatment with high persistence and low incentives than in the treatment with low persistence and high incentives.*

6 Discussion

The ways in which habits are conceived vary across fields. In macroeconomics, habits are preference-based; they are identified with non-separable utilities. Psychologists and neuroscientists emphasize procedural aspects of habitual behavior: it is automatic, subconscious

[9] We have excluded the interaction term $\delta\theta_n^1$, since the state realizations satisfied $\theta_1^n = \theta_2^n$ for all n in CNS.

and fast, unlike the conscious decision process. Another definitory aspect of habits within neuroscience is behavior-based: habits are triggered by historically formed cues and they may continue to be employed beyond the span of their functionality (e.g. Dezfouli & Balleine (2012)). Our definition of habits falls within the last, behavior-based approach; we say that habit is formed if the first-period variables irrelevant for the second-period payoffs predict the second-period behavior.

Models rooted in neuroscience and psychology, such as that of Camerer et al. (n.d.) and Laibson (2001) focus on the procedural aspects of habitual behavior. Camerer et al. specify an automatic process which governs switches between the habitual (often called model-free) mode and preference-based mode of optimization (often called model-based mode). The habitual mode is typically modeled as a reinforcement-learning procedure, such as that of Roth & Erev (1995), in which historically successful actions are automatically reinforced in a reward-contingent loop.

We differ in several aspects from such reinforcement learning-based concepts of habits. First, we do not attempt to study the dynamics of habitualization. The Markov process of payoff states and action consequences are extremely simple in our experimental design and hence our subjects do not need to learn them. Second, unlike in the reinforcement-learning models, our DM is not backward-looking. Rather, she is likely to repeat her first-period action in some of our treatments, because she chose it based on a first-period signal, and this signal correlates with the optimal second-period action. Third, unlike the reinforcement-learning models, our model is based on optimization of the habit strength, and this optimization generates comparative-statics predictions.

The reinforcement-learning models of habits assume a causal link between the past and the current actions; the DM repeats past actions because they are successful. While the past and current actions become correlated in our framework (in some treatments), the connection is not causal: the second-period choice of our DM is influenced by her first-period signal rather than by her first-period choice. Yet, since the first-period choice and signal are perfectly correlated, we predict serial correlation in the action sequence. However, should an intervention revert the agent's first-period choice after her first-period signal was acquired, the second-period action would be unaffected in our model.

In summary, our experiment reveals a significant degree of sophistication in habit formation, cue selection, and in the choice of habit strength. This sophistication favors concepts of habits based on explicit optimization under information frictions relative to models based on automatic rules. Yet, our data also indirectly document myopia in information acquisition–a feature inconsistent with rational, forward-looking information-acquisition models.

A Appendix

A.1 Proofs

Proof of Lemma 1. We solve the problem backwards. Let p_2 be the probability that the DM assigns to $\theta_2 = 1$ at the beginning of the period 2. Let $v(q) = s \max\{q, 1-q\}$ be the DM's expected optimized second-period payoff at a given value of the posterior $q = \Pr(\theta_2 = 1 \mid x_1, y, x_2)$. The DM chooses a random second-period posterior $\hat{q}_2$ that maximizes the second-period payoff net of the second-period information cost:

$$\begin{aligned} \max_{\hat{q}_2} \quad & \mathrm{E}_{\hat{q}_2}\left[v(\hat{q}_2) - H(p_2) + H(\hat{q}_2)\right] \\ \text{s.t.:} \quad & \mathrm{E}_{\hat{q}_2}\, \hat{q}_2 = p_2. \end{aligned}$$

We let $w_2(p_2)$ be the value of the above problem. This second-period problem can be solved by the concavification of $v(\cdot) + H(\cdot)$; see Caplin & Dean (2013). Since θ_2 has binary support, the support of the second-period posterior $\hat{q}_2$ is at most binary. Additionally, due to our regularity condition, both $a_2 = 0$ and $a_2 = 1$ occur with positive probabilities, and thus the support of $\hat{q}_2$ contains one value $\underline{q}$ below and one value $\overline{q}$ above 1/2. The values $\underline{q}$ and $\overline{q}$ are independent of p_2.

In the setting without feedback, p_2 attains the random value $\hat{p}_2 = \gamma\hat{q}_1 + (1-\gamma)(1-\hat{q}_1)$ where $\hat{q}_1 = \Pr(\theta_1 \mid x_1)$ is the random posterior at the end of the period 1. Let

$$v_1^n(q_1) = v(q_1) + w_2\left(\gamma q_1 + (1-\gamma)(1-q_1)\right)$$

be the value of the DM whose first-period posterior belief $\hat{q}_1$ attains value q_1; it consists of the immediate expected payoff $v(q_1)$ for the first-period action and of the continuation payoff $w_2\left(\gamma q_1 + (1-\gamma)(1-q_1)\right)$. The DM chooses a random first-period posterior $\hat{q}_1$ that solves

$$\begin{aligned} \max_{\hat{q}_1} \quad & \mathrm{E}_{\hat{q}_1}\left[v_1^n(\hat{q}_1) - H(p_1) + H(\hat{q}_1)\right] \\ \text{s.t.:} \quad & \mathrm{E}_{\hat{q}_1}\, \hat{q}_1 = p_1, \end{aligned}$$

where p_1 is the prior probability that $\theta_1 = 1$. This problem can also be solved by concavification. Again, due to the binary support of θ_1 and the regularity condition, $\hat{q}_1$ attains two values that we denote by $\underline{q}^n < 1/2$ and $\overline{q}^n > 1/2$. Again, this support is independent of p_1.

In the setting with feedback, p_2 attains the random value $\hat{p}_2 = \gamma y + (1-\gamma)(1-y)$ and

the DM solves

$$\begin{array}{ll} \max\limits_{\hat{q}_1} & \mathrm{E}_{\hat{q}_1}\left[v_1^f(\hat{q}_1) - H(p_1) + H(\hat{q}_1)\right] \\ \text{s.t.:} & \mathrm{E}_{\hat{q}_1}\,\hat{q}_1 = p_1, \end{array}$$

where

$$v_1^f(q_1) = v(q_1) + \mathrm{E}_y\left[w_2\left(\gamma y + (1-\gamma)(1-y)\right)\right].$$

Since the continuation value $w_2\left(\gamma y + (1-\gamma)(1-y)\right)$ does not depend on the first-period posterior, the DM chooses the support of $\hat{q}_1$ to be the same as that of $\hat{q}_2$; $\hat{q}_1$ attains values $\underline{q}$ and $\overline{q}$, again, independently of p_1.

The joint distribution $\pi(\theta_1, \theta_2, a_1, a_2)$ is uniquely determined by the binary supports of $\hat{q}_1$ and $\hat{q}_2$. □

Proof of Proposition 1. Statement 1.: The support of the random second-period posterior is independent of the first-period posterior and of y. Since θ_1 and θ_2 are independent, the prior p_2 at the beginning of period 2 is independent of θ_1 and a_1. Thus, the random second-period posterior (and hence a_2) is a r.v. independent of θ_1 and a_1, conditionally on θ_2, as needed.

Statement 2.:

$$\pi(a_2 = 1 \mid \theta_2, a_1 = 1, \theta_1) = \Pr\left(\hat{q}_2\left(\hat{q}_1\right) = \overline{q} \mid \theta_2, \hat{q}_1 = \overline{q}^n\right),$$

$$\pi(a_2 = 1 \mid \theta_2, a_1 = 0, \theta_1) = \Pr\left(\hat{q}_2\left(\hat{q}_1\right) = \overline{q} \mid \theta_2, \hat{q}_1 = \underline{q}^n\right).$$

The right-hand sides do not depend on θ_1, as needed. It suffices to prove that for each $\theta_2 \in \{0, 1\}$, the first expression exceeds the latter. We consider the case $\theta_2 = 1$; the computation for $\theta_2 = 0$ is analogous.

$$\begin{aligned} \pi(a_2 = 1 \mid \theta_2 = 1, a_1 = 1) &= \Pr\left(\hat{q}_2\left(\hat{q}_1\right) = \overline{q} \mid \theta_2 = 1, \hat{q}_1 = \overline{q}^n\right) \\ &= \Pr\left(\theta_2 = 1 \mid \hat{q}_2(\hat{q}_1) = \overline{q}, \hat{q}_1 = \overline{q}^n\right) \frac{\Pr\left(\hat{q}_2(\hat{q}_1) = \overline{q} \mid \hat{q}_1 = \overline{q}^n\right)}{\Pr\left(\theta_2 = 1 \mid \hat{q}_1 = \overline{q}^n\right)} \\ &= \overline{q}\frac{\frac{\overline{q}^n\gamma + (1-\overline{q}^n)(1-\gamma) - \underline{q}}{\overline{q} - \underline{q}}}{\overline{q}^n\gamma + (1-\overline{q}^n)(1-\gamma)} \\ &= \frac{\overline{q}}{\overline{q} - \underline{q}}\frac{\overline{q}^n\gamma + (1-\overline{q}^n)(1-\gamma) - \underline{q}}{\overline{q}^n\gamma + (1-\overline{q}^n)(1-\gamma)} \\ &= \varphi\left(\overline{q}^n\gamma + (1-\overline{q}^n)(1-\gamma)\right), \end{aligned}$$

where $\varphi(p_2) = \frac{\overline{q}}{\overline{q}-\underline{q}}\frac{p_2 - \underline{q}}{p_2}$. An analogous computation implies that $\pi(a_2 = 1 \mid \theta_2 = 1, a_1 = 0) = \varphi\big(\underline{q}^n\gamma + (1-\underline{q}^n)(1-\gamma)\big)$. The claim follows from the monotonicity of φ.

Statement 3.: When the DM receives feedback, then her belief at the beginning of period 2 is $\hat{p}_2 = \gamma\theta_1 + (1-\gamma)(1-\theta_1)$. Since $\hat{p}_2$ and the values of the second-period posteriors $\overline{q}$ and $\underline{q}$ do not depend on a_1, $\pi(a_2 \mid \theta_2, \theta_1, a_1)$ does not depend on a_1, as needed.

Let us consider $\theta_2 = 1$ (the case of $\theta_2 = 0$ is again analogous). The values of the second-period posteriors, $\overline{q}$ and $\underline{q}$, are the same as in the setting without feedback. Thus again, as in the proof of Statement 2,

$$\pi(a_2 = 1 \mid \theta_2 = 1, \theta_1 = 1) = \varphi(\gamma) > \varphi(1-\gamma) = \pi(a_2 = 1 \mid \theta_2 = 1, \theta_1 = 0).$$

□

The next lemma is an auxiliary result that we use in the proof of Proposition 3. It characterizes the habit strength as a function of the posterior values. To economize on notation, we write from now on $q_2 \in [1/2, 1) = \pi(\theta_2 = 1 \mid a_2 = 1) = \overline{q}$ for the higher of the two realizations of the second-period posterior, and we note that thanks to the symmetry of the setting, $\pi(\theta_2 = 1 \mid a_2 = 0) = 1 - q_2$ both in the settings with and without feedback. Similarly, we write $q_1 \in [1/2, 1)$ for the high realization of the first-period posterior $\pi(\theta_1 = 1 \mid a_1 = 1)$ and note that $\pi(\theta_1 = 1 \mid a_1 = 0) = 1 - q_1$ in the both settings. We recall that the value of q_1 depends on the feedback specification. Finally, in the setting without feedback, we let p_2 stand for the belief at the beginning of period 2 of the DM who chose $a_1 = 1$ in period 1. It is $p_2 = \pi(\theta_2 = 1 \mid a_1 = 1) = \gamma q_1 + (1-\gamma)(1-q_1)$. We recall that the belief at the beginning of period 2 is $\gamma\theta_1 + (1-\gamma)(1-\theta_1)$ in the setting with feedback.

Lemma 2. *1. In the setting without feedback, the habit strength is*

$$\phi_{a_1}(\theta_2) = \frac{p_2 + q_2 - 1}{q_2 - p_2} \frac{1 - p_2}{p_2}. \tag{2}$$

2. In the setting with feedback, the habit strength is

$$\phi_{\theta_1}(\theta_2) = \frac{\gamma + q_2 - 1}{q_2 - \gamma} \frac{1 - \gamma}{\gamma}. \tag{3}$$

(Observe that habit strength is independent of θ_2 in both cases.)

Proof of Lemma 2. Statement 1.: By its definition, habit strength when the cue is the first-

period action is,

$$\begin{aligned}\phi_{a_1}(\theta) &= \frac{\pi(a_2=\theta \mid \theta_2=\theta, a_1=\theta)}{\pi(a_2=\theta \mid \theta_2=\theta, a_1=1-\theta)} \\ &= \frac{\pi(\theta_2=\theta \mid a_2=\theta, a_1=\theta)\pi(a_2=\theta \mid a_1=\theta)/\pi(\theta_2=\theta \mid a_1=\theta)}{\pi(\theta_2=\theta \mid a_2=\theta, a_1=1-\theta)\pi(a_2=\theta \mid a_1=1-\theta)/\pi(\theta_2=\theta \mid a_1=1-\theta)}.\end{aligned}$$

Since the posterior $\pi(\theta_2=\theta \mid a_2=\theta, a_1)$ is independent of a_1,

$$\begin{aligned}\phi_{a_1}(\theta) &= \frac{\pi(a_2=\theta \mid a_1=\theta)/\pi(\theta_2=\theta \mid a_1=\theta)}{\pi(a_2=\theta \mid a_1=1-\theta)/\pi(\theta_2=\theta \mid a_1=1-\theta)} \\ &= \frac{(p_2+q_2-1)/p_2}{(q_2-p_2)/(1-p_2)},\end{aligned}$$

where we have used the martingale condition imposed on the second-period posteriors to derive $\pi(a_2=\theta \mid a_1=\theta)=\frac{p_2+q_2-1}{2q_2-1}$ and $\pi(a_2=\theta \mid a_1=1-\theta)=\frac{q_2-p_2}{2q_2-1}$, and we have noted that $\pi(\theta_2=\theta \mid a_1=\theta)=p_2$ and $\pi(\theta_2=\theta \mid a_1=1-\theta)=1-p_2$.

Statement 2.: When the cue is the first-period state then the habit strength is defined as

$$\begin{aligned}\phi_{\theta_1}(\theta) &= \frac{\pi(a_2=\theta \mid \theta_2=\theta, \theta_1=\theta)}{\pi(a_2=\theta \mid \theta_2=\theta, \theta_1=1-\theta)} \\ &= \frac{\pi(\theta_2=\theta \mid a_2=\theta, \theta_1=\theta)\pi(a_2=\theta \mid \theta_1=\theta)/\pi(\theta_2=\theta \mid \theta_1=\theta)}{\pi(\theta_2=\theta \mid a_2=\theta, \theta_1=1-\theta)\pi(a_2=\theta \mid \theta_1=1-\theta)/\pi(\theta_2=\theta \mid \theta_1=1-\theta)}.\end{aligned}$$

Since the posterior $\pi(\theta_2=\theta \mid a_2=\theta, \theta_1)$ is independent of θ_1,

$$\begin{aligned}\phi_{\theta_1}(\theta) &= \frac{\pi(a_2=\theta \mid \theta_1=\theta)/\pi(\theta_2=\theta \mid \theta_1=\theta)}{\pi(a_2=\theta \mid \theta_1=1-\theta)/\pi(\theta_2=\theta \mid \theta_1=1-\theta)} \\ &= \frac{(\gamma+q_2-1)/\gamma}{(q_2-\gamma)/(1-\gamma)},\end{aligned}$$

where we have used the martingale condition imposed on the second-period posteriors to derive that $\pi(a_2=\theta \mid \theta_1=\theta)=\frac{\gamma+q_2-1}{2q_2-1}$ and $\pi(a_2=\theta \mid \theta_1=1-\theta)=\frac{q_2-\gamma}{2q_2-1}$. □

Proof of Propositions 2 and 3. Setting without feedback: Since the cost function is symmetric, the DM solves

$$\max_{q_1,q_2}\left\{s\left(q_1+q_2\right)-H(p_1)+H(q_1)-H\left(\gamma q_1+(1-\gamma)(1-q_2)\right)+H(q_2)\right\}.$$

The second-period posterior $q_2 \geq 1/2$ solves the first-order condition

$$s = -H'(q_2). \tag{4}$$

Because $-H'(q_2)$ increases on $[1/2, 1)$ and attains values in $[0, \infty)$, (4) has a unique solution $q_2 \in (1/2, 1)$ that increases with the stake s.

Similarly, letting

$$\tilde{H}(q) = H(q) - H(\gamma q + (1-\gamma)(1-q)),$$

$q_1 \geq 1/2$ solves

$$s = -\tilde{H}'(q_1). \tag{5}$$

By part A.4 of Assumption A,

$$\tilde{H}''(q) = H''(q) - (2\gamma - 1)^2 H''(\gamma q + (1-\gamma)(1-q)) < 0.$$

Thus, $-\tilde{H}'(q_1)$ increases and attains values in $[0, \infty)$ for $q \in [1/2, 1)$. Thus also (5) has a unique solution $q_1 \in (1/2, 1)$ that increases with s. Additionally, $q_2 < q_1$ because $-H'(q) > -\tilde{H}'(q)$ for all $q \in [1/2, 1)$; this proves Proposition 2.

To prove Proposition 3, we first analyze comparative statics with respect to γ. Posterior q_2 is independent of γ. Posterior q_1 increases with γ since $-\tilde{H}'(q)$ decreases with γ for all $q \in [1/2, 1)$. Hence $p_2 = \gamma q_1 + (1-\gamma)(1-q_1)$ increases with γ too. We notice from (2) that $\frac{\partial \phi_{a_1}}{\partial p_2} > 0$. Thus, ϕ_{a_1} increases in γ, as needed.

We now examine comparative statics with respect to s under the assumption that $H(\cdot)$ is Shannon entropy. We combine (4) and (5) to get

$$H'(q_2) = H'(q_1) - (2\gamma - 1) H'\left(\gamma q_1 + (1-\gamma)(1-q_1)\right),$$

and we express q_2 as an increasing function of q_1. Using that $H'(q) = \log \frac{1-q}{q}$, we get

$$q_2(q_1) = \frac{1}{1+x},$$

with $x = \frac{1-q_1}{q_1} \left(\frac{p_2}{1-p_2}\right)^{2\gamma-1} > 0$, where we remind that $p_2 = \gamma q_1 + (1-\gamma)(1-q_1)$ is a function

of q_1. We prove that $\frac{d^2q_2}{dq_1^2}$ is positive.

$$\begin{aligned}
\frac{dq_2}{dq_1} &= \frac{-1}{(1+x)^2}\frac{dx}{dq_1} \\
&= \frac{-1}{(1+x)^2}\left(\frac{-1}{q_1^2}\left(\frac{p_2}{1-p_2}\right)^{2\gamma-1} + \frac{1-q_1}{q_1}\left(\frac{p_2}{1-p_2}\right)^{2\gamma-2}\frac{(2\gamma-1)^2}{(1-p_2)^2}\right) \\
&= \frac{-1}{(1+x)^2}\frac{1-q_1}{q_1}\left(\frac{p_2}{1-p_2}\right)^{2\gamma-1}\left(\frac{-1}{q_1(1-q_1)} + \frac{(2\gamma-1)^2}{p_2(1-p_2)}\right) \\
&= \frac{x}{(1+x)^2}\left(\frac{1}{q_1(1-q_1)} - \frac{(2\gamma-1)^2}{p_2(1-p_2)}\right).
\end{aligned}$$

$$\begin{aligned}
\frac{d^2q_2}{dq_1^2} &= \frac{(1-x)}{(1+x)^3}\frac{dx}{dq_1}\left(\frac{1}{q_1(1-q_1)} - \frac{(2\gamma-1)^2}{p_2(1-p_2)}\right) + \frac{x}{(1+x)^2}\left(\frac{2q_1-1}{q_1^2(1-q_1)^2} - (2\gamma-1)^3\frac{2p_2-1}{p_2^2(1-p_2)^2}\right) \\
&= -\frac{(1-x)x}{(1+x)^3}\left(\frac{1}{q_1(1-q_1)} - \frac{(2\gamma-1)^2}{p_2(1-p_2)}\right)^2 + \frac{x}{(1+x)^2}\left(\frac{2q_1-1}{q_1^2(1-q_1)^2} - (2\gamma-1)^3\frac{2p_2-1}{p_2^2(1-p_2)^2}\right) \\
&= -\frac{(1-x)x}{(1+x)^3}\left(\frac{1}{q_1(1-q_1)} - \frac{(2\gamma-1)^2}{p_2(1-p_2)}\right)^2 + \frac{x}{(1+x)^2}\left(\frac{2q_1-1}{q_1^2(1-q_1)^2} - (2\gamma-1)^4\frac{2q_1-1}{p_2^2(1-p_2)^2}\right) \\
&> \frac{x(2q_1-1)}{(1+x)^2}\left(-\left(\frac{1}{q_1(1-q_1)} - \frac{(2\gamma-1)^2}{p_2(1-p_2)}\right)^2 + \left(\frac{1}{q_1^2(1-q_1)^2} - \frac{(2\gamma-1)^4}{p_2^2(1-p_2)^2}\right)\right) \\
&= \frac{x(2q_1-1)}{(1+x)^2}\left(\frac{1}{q_1(1-q_1)} - \frac{(2\gamma-1)^2}{p_2(1-p_2)}\right)\frac{2(2\gamma-1)^2}{p_2(1-p_2)} \\
&> 0,
\end{aligned}$$

where we have used $(2\gamma-1)(2q_1-1) = (2p_2-1)$ in the third step, $\frac{1-x}{1+x} = 2q_2 - 1 < (2q_1-1)$ in the fourth step, and $q_1 > p_2$ in the last step.

We notice from (2) that ϕ_{a_1} decreases with s if $\frac{p_2+q_2-1}{q_2-p_2}$ decreases with s (since $\frac{1-p_2}{p_2}$

decreases with q_1 and hence with s). Thus, ϕ_{a_1} decreases with s if

$$\begin{aligned}
0 &> \frac{d}{ds}\frac{p_2+q_2-1}{q_2-p_2} \\
&= \frac{(2q_2-1)\frac{dp_2}{ds}-(2p_2-1)\frac{dq_2}{ds}}{(q_2-p_2)^2} \\
&= \left((2q_2-1)\frac{dp_2}{dq_1}-(2p_2-1)\frac{dq_2}{dq_1}\right)\frac{\frac{dq_1}{ds}}{(q_2-p_2)^2} \\
&= \left((2q_2-1)(2\gamma-1)-(2p_2-1)\frac{dq_2}{dq_1}\right)\frac{\frac{dq_1}{ds}}{(q_2-p_2)^2} \\
&= \left(2q_2-1-(2q_1-1)\frac{dq_2}{dq_1}\right)\frac{(2\gamma-1)\frac{dq_1}{ds}}{(q_2-p_2)^2},
\end{aligned}$$

where we have used $\frac{dp_2}{dq_1}=(2\gamma-1)$ for the third equality, and $\frac{2p_2-1}{2\gamma-1}=2q_1-1$ to establish the fourth equality. Therefore, it suffices to prove that

$$2q_2-1<(2q_1-1)\frac{dq_2}{dq_1}.$$

We observe that $q_2=1/2$ when $q_1=1/2$. Thus, by the Mean value theorem, there exists $1/2<\tilde{q}_1<q_1$ such that,

$$2q_2-1=(2q_1-1)\left.\frac{dq_2}{dq_1}\right|_{\tilde{q}_1}<(2q_1-1)\left.\frac{dq_2}{dq_1}\right|_{q_1},$$

where the inequality follows from the fact that $\frac{dq_2}{dq_1}$ increases with q_1.

Setting with feedback: Again, q_2 solves (4) and thus q_2 increases with s and it is independent of γ. It follows from (3) that ϕ_{θ_1} increases with γ. Finally, ϕ_{θ_1} decreases with q_2, and hence with s. □

	Frequency of $a_1 = \theta_1$	Frequency of $a_2 = \theta_2$	Frequency of $a_2 = a_1$	Frequency of $a_1 = \theta_2$
IF	.73	.66	.54	.55
IN	.79	.79	.52	.49
CF	.82	.81	.73	.70
CN	.79	.74	.79	.70

Table 5: Preliminary session: Data summary

	IF	IN	CF	CN
a_1	-.029 (.091) .748	.037 (.056) .501	.037 (.055) .500	**.449** (**.083**) .000
θ_1	.129 (.101) .203	-.023 (.046) .617	**.408** (**.203**) .000	$-$.095 (.065) .144
θ_2	**.267** (**.065**) .000	**.552** (**.072**) .000	**.317** (**.082**) .000	**.436** (**.099**) .000

Table 6: Preliminary session: Average marginal treatment effects, (their standard errors), and the p-values in the second lines. Bold values indicate significance at the 1% level.

A.2 Preliminary Session

We ran a preliminary session prior to the regular sessions. Sixteen participating subjects obtained a \$15 show-up fee and an additional \$5 for a correct answer to the counting task (randomly selected at the end of the experiment). The parameters were: $\gamma = .5$ in treatments with independent states (I) and $\gamma = .75$ in treatments with correlated states (C). As in the regular sessions, θ_1 was revealed in between periods in the treatments F with feedback and it was not revealed in treatments N without feedback. The treatment order was IF, CF, IN, CN.

The basic data description in Table 5 and the estimated average marginal treatment effects in Table 6 are consistent with the results from the regular sessions. However, in this session, the subjects were free to leave immediately once they finished all their counting tasks in the last treatment (CN), which affected their information processing costs in an uncontrolled manner, and thus we omit the pilot data from the main analysis.

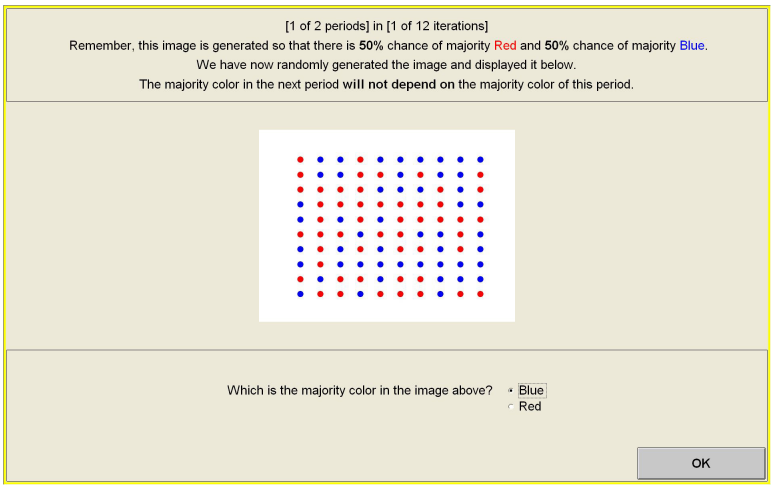

Figure 1: A scaled-down screenshot of the counting task.

A.3 Experimental instructions

Instructions

Welcome to the experiment! Please take a record sheet at the front if you don't have one already. Please do not use the computers during the instructions. When it is time to use the computer, please follow the instructions precisely.(Repeat if necessary.)

Please raise your hand if you need a pencil. Please put away and silence all your personal belongings, especially your phone. We need your full attention during the experiment.

Raise your hand at any point if you cannot see or hear well.

The experiment you will be participating in today is an experiment in decision making. At the end of the experiment, you will be paid for your participation in cash. The amount you earn depends on your decisions and on chance. You will be using the computer for the experiment, and all decisions will be made through the computer. DO NOT socialize or talk during the experiment.

All instructions and descriptions that you will be given in the experiment are accurate and true. In accordance with the policy of this lab, at no point will we attempt to deceive you in any way.

If you have any questions, raise your hand and your question will be answered out loud so everyone can hear.

After you have completed all the tasks, please wait while everyone else finishes his or her tasks. Once everyone has completed the experiment, I will ask you to fill in the questionnaire. After the questionnaire you will collect your earnings and leave.

//

I will now describe the main features of the experiment and show you how to use the software. Again, if you have any questions during this period, please raise your hand.

You will be presented with a series of choices to make. There will be four SETS of choices in today's experiment. Each set contains twelve ITERATIONS, and each iteration has two PERIODS. In each period, you will be shown a picture of 100 dots. Each dot will be either RED or BLUE. We have displayed an example of such a screen on your computer monitor. (show an example screen)

This is an example of the screens you will see during the experiment. In every period, the picture will contain either 51 red dots and 49 blue dots, or instead, 51 blue dots and 49 red dots. We will call these two cases MAJORITY RED and MAJORITY BLUE, respectively. In each case, the dots are randomly allocated to the positions in the matrix. In each period the computer will choose randomly between MAJORITY RED and MAJORITY BLUE. You will be told in advance how likely each case is to happen.

In each period, you will be asked to determine if the image is MAJORITY RED or MAJORITY BLUE. While you may take as much time as you need to make your choice, the image will disappear after 45 seconds.

I am now going to describe the details of the experiment.

The experiment is divided into four SETS. In each set, you will be presented with twelve iterations, and each iteration consists of two periods, each with its own image. The rules for the 12 iterations within each set are identical, but the rules are different in different sets.

In PERIOD 1 of each iteration, the image is always generated so that there is an equal chance of MAJORITY RED and MAJORITY BLUE, meaning that there is a 50% chance of MAJORITY RED and a 50% chance of MAJORITY BLUE.

In period 2 of each iteration, the image will be generated in a way that differs across sets. In some sets, the majority color for period 2 is chosen in a way that is completely separate from the period 1 image, and is randomly generates so that there is an equal chance of MAJORITY RED and MAJORITY BLUE, just like the period 1 image. But in other sets, the period 2 image depends on the majority color of the period 1 image. In these sets, the computer generates the period 2 image so that there is a 75% chance that the majority color matches the period 1 majority color, and a 25% chance that the majority color is different from the period 1 majority color.

It is important to remember that while the periods within each iteration may be related to each other, the periods across iterations are never related.

After making your choices, you will always be told what the majority color was, but the timing of this differs from set to set. In some sets, the majority colors will be revealed after every period. In other sets, the majority colors for an iteration will not be revealed until you complete both periods. Before each set, you will be told about the timing of the feedback you will receive.

The amount of money you will receive at the end of the experiment depends on your choices. After we have completed all four sets, you will have made 96 choices (4 sets times 12 iterations times 2 periods). The computer software will randomly select one of these 96 periods. Your payment will be determined by your choice in that single period. If your choices in the randomly chosen period matches the majority color, you will earn an additional $5 dollars on top of the $15 show-up fee. Otherwise, you will receive no additional payment, but you will still receive the show-up fee.

After you complete the last set, please wait until we start the questionnaire part. After you finish the questionnaire, please fill your record sheet on the desk. I will pay one by one to keep everyone's privacy.

To summarize, remember that we have four sets in the experiment today. Each set consists of 12 iterations, and each iteration consists of two periods. The sets will vary in how likely it is that the majority colors are the same for both periods within an iteration, and in the timing that the majority colors are revealed. Please raise your hand if you have any questions.

(1) FI/FC/NI/NC

Feedback/IID:

In the next set of twelve iterations, the majority color for period 2 is randomly generated so that there is an equal chance of MAJORITY RED and MAJORITY BLUE, and it does not depend on the majority color in the first period.

The majority colors will be revealed after every period, so that you will be told the majority color from period 1 before you see the image for period 2. Please raise your hand if you have any question.

Feedback/Corr.:

In the next set of twelve iterations, the majority color for period 2 is randomly generated so that there is a 75% chance that the majority color matches the majority color from period 1, and a 25% chance that the majority color is different from period 1.

The majority colors will be revealed after every period, so that you will be told the majority color from period 1 before you see the image for period 2. Please raise your hand if you have any question.

No Feedback/IID:

In the next set of twelve iterations, the majority color for period 2 is randomly generated so that there is an equal chance of MAJORITY RED and MAJORITY BLUE, and it does not depend on the majority color in the first period.

The majority colors for both periods of an iteration will be revealed only at the end of each iteration, so that you will see the period 2 image before being told the majority color from period 1. Please raise your hand if you have any question.

No Feedback/Corr.:

In the next set of twelve iterations, the majority color for period 2 is randomly generated so that there is a 75% chance that the majority color matches the period 1 majority color, and a 25% chance that the majority color is different from the period 1 majority color.

The majority colors for both periods of an iteration will be revealed only at the end of each iteration, so that you will see the period 2 image before being told the majority color from period 1. Please raise your hand if you have any question.

(2) NI/NC/FI/FC

No Feedback/IID:

In the next set of twelve iterations, the majority color for period 2 is randomly generated so that there is an equal chance of MAJORITY RED and MAJORITY BLUE, and it does not depend on the majority color in the first period.

The majority colors for both periods of an iteration will be revealed only at the end of each iteration, so that you will see the period 2 image before being told the majority color from period 1. Please raise your hand if you have any question.

No Feedback/Corr.:

In the next set of twelve iterations, the majority color for period 2 is randomly generated so that there is a 75% chance that the majority color matches the period 1 majority color, and a 25% chance that the majority color is different from the period 1 majority color.

The majority colors for both periods of an iteration will be revealed only at the end of each iteration, so that you will see the period 2 image before being told the majority color from period 1. Please raise your hand if you have any question.

Feedback/IID:

In the next set of twelve iterations, the majority color for period 2 is randomly generated so that there is an equal chance of MAJORITY RED and MAJORITY BLUE, and it does not depend on the majority color in the first period.

The majority colors will be revealed after every period, so that you will be told the majority color from period 1 before you see the image for period 2. Please raise your hand if you have any question.

Feedback/Corr.:

In the next set of twelve iterations, the majority color for period 2 is randomly generated so that there is a 75% chance that the majority color matches the majority color from period 1, and a 25% chance that the majority color is different from period 1.

The majority colors will be revealed after every period, so that you will be told the majority color from period 1 before you see the image for period 2. Please raise your hand if you have any question.

References

Abel, A. (1990), 'Asset prices under habit formation and catching up with the Joneses', *American Economic Review* **80**(2), 38–42.

Andrews, B. R. (1903), 'Habit', *The American Journal of Psychology* **14**(2), 121–149.

Angeletos, G.-M. & Huo, Z. (2018), Collective myopia and habit, *in* '2018 Meeting Papers', number 236, Society for Economic Dynamics.

Camerer, C., Landry, P. & Webb, R. (n.d.), The neuroeconomics of habit, *in* A. Kirman & M. Tedeschi, eds, 'The State of Mind', Oxford University Press.

Caplin, A. & Dean, M. (2013), Behavioral implications of rational inattention with Shannon entropy, Technical report, National Bureau of Economic Research Working Paper 19318.

Caplin, A. & Dean, M. (2015), 'Revealed preference, rational inattention, and costly information acquisition', *American Economic Review* **105**(7), 2183–2203.

Caplin, A., Dean, M. & Leahy, J. (2017), 'Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy', *National Bureau of Economic Research Working Paper Series* **No. 23652**.

Dezfouli, A. & Balleine, B. W. (2012), 'Habits, action sequences and reinforcement learning', *European Journal of Neuroscience* **35**(7), 1036–1051.

Fischbacher, U. (2007), 'z-tree: Zurich toolbox for ready-made economic experiments', *Experimental economics* **10**(2), 171–178.

Khaw, M. W., Stevens, L. & Woodford, M. (2017), 'Discrete adjustment to a changing environment: Experimental evidence', *Journal of Monetary Economics* **91**, 88–103.

Khaw, M. W. & Zorrilla, O. (2018), 'Deeper habits'.

Laibson, D. (2001), 'A cue-theory of consumption', *The Quarterly Journal of Economics* **116**(1), 81–119.

Lally, P., Van Jaarsveld, C. H., Potts, H. W. & Wardle, J. (2010), 'How are habits formed: Modelling habit formation in the real world', *European Journal of Social Psychology* **40**(6), 998–1009.

Mackowiak, B. & Wiederholt, M. (2009), 'Optimal sticky prices under rational inattention', *American Economic Review* **99**(3), 769–803.

Matějka, F. & McKay, A. (2015), 'Rational inattention to discrete choices: A new foundation for the multinomial logit model', *American Economic Review* **105**(1), 272–98.

McNemar, Q. (1947), 'Note on the sampling error of the difference between correlated proportions or percentages', *Psychometrika* **12**(2), 153–157.

Morris, S. & Strack, P. (2017), 'The Wald Problem and the Equivalence of Sequential Sampling and Static Information Costs', *SSRN Electronic Journal* .

Pollak, R. A. (1970), 'Habit formation and dynamic demand functions', *Journal of Political Economy* **78**(4, Part 1), 745–763.

Roth, A. E. & Erev, I. (1995), 'Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term', *Games and Economic Behavior* **8**(1), 164–212.

Schmitt-Grohé, S. & Uribe, M. (2007), 'Habit persistence', *The New Palgrave Dictionary of Economics, second edition, Palgrave Macmillan, NY, NY* .

Sims, C. A. (2003), 'Implications of rational inattention', *Journal of Monetary Economics* **50**(3), 665–690.

Steiner, J., Stewart, C. & Matějka, F. (2017), 'Rational inattention dynamics: Inertia and delay in decision-making', *Econometrica* **85**(2), 521–553.

Wood, W. & Neal, D. T. (2007), 'A new look at habits and the habit-goal interface.', *Psychological Review* **114**(4), 843.