Cognitive Psychology 93 (2017) 18–43

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

Learning in settings with partial feedback and the wavy recency effect of rare events

Ori Plonsky [a,*], Ido Erev [a,b]

[a] Max Wertheimer Minerva Center for Cognitive Studies, Faculty of Industrial Engineering and Management, Technion, Israel
[b] Warwick Business School, Warwick University, UK

ARTICLE INFO

Article history:
Accepted 15 January 2017

Keywords:
Law of effect
Decisions from experience
Negative recency
The gambler's fallacy
Contingencies of reinforcements
Pattern learning

ABSTRACT

Analyses of human learning reveal a discrepancy between the long- and the short-term effects of outcomes on subsequent choice. The long-term effect is simple: favorable outcomes increase the choice rate of an alternative whereas unfavorable outcomes decrease it. The short-term effects are more complex. Favorable outcomes can decrease the choice rate of the best option. This pattern violates the positive recency assumption that underlies the popular models of learning. The current research tries to clarify the implications of these results. Analysis of wide sets of learning experiments shows that rare positive outcomes have a wavy recency effect. The probability of risky choice after a successful outcome from risk-taking at trial t is initially (at $t+1$) relatively high, falls to a minimum at $t+2$, then increases for about 15 trials, and then decreases again. Rare negative outcomes trigger a wavy reaction when the feedback is complete, but not under partial feedback. The difference between the effects of rare positive and rare negative outcomes and between full and partial feedback settings can be described as a reflection of an interaction of an effort to discover patterns with two other features of human learning: surprise-triggers-change and the hot stove effect. A similarity-based descriptive model is shown to capture well all these interacting phenomena. In addition, the model outperforms the leading models in capturing the outcomes of data used in the 2010 Technion Prediction Tournament.

1. Introduction

The "law of effect" (Thorndike, 1898) is perhaps the most important and robust finding in the learning literature. It suggests that in a given situation, good outcomes increase the probability of the reinforced behavior while bad outcomes decrease it. Consider for example a repeated choice task with many trials, and specifically, the effect of an outcome at some trial t on all subsequent choices. According to the common interpretation of the law of effect, at all trials following trial t, the choice rate of an alternative that generates a favorable outcome at trial t should, ceteris paribus, be at least as high as that alternative's choice rate had that outcome been unfavorable. More specifically, the attractiveness of an alternative is

Abbreviations: CA, contingent average; TPT, Technion Prediction Tournament; CAT, Contingent Average and Trend; CATIE, Contingent Average, Trend, Inertia, and Exploration; CAB-*k*, contingent average based on *k* outcomes.

* Corresponding author at: Max Wertheimer Minerva Center for Cognitive Studies, Faculty of Industrial Engineering and Management, Technion, Haifa 3200003, Israel.

E-mail address: plonsky@campus.technion.ac.il (O. Plonsky).

http://dx.doi.org/10.1016/j.cogpsych.2017.01.002
0010-0285/

predicted to increase immediately after the alternative generates a good outcome, and this effect should gradually fade in time (as more outcomes, good and bad, are generated by the possible alternatives). This interpretation of the law of effect predicts outcomes have a positive recency effect on choice.

Although in the long term the common interpretation of the law of effect summarizes well the effects of outcomes on choice, analyses of human learning reveal that the short term effects are more complex. Specifically, following runs of identical outcomes, decision makers often behave as if they expect a change to occur ("the gabmler's fallacy"; Jarvik, 1951), and following short stretches of alternating outcomes, decision makers tend to expect the alternation to continue (Anderson, 1960). More recently, Plonsky, Teodorescu, and Erev (2015) show rare events have an even more complex effect on subsequent choice: a *wavy recency* effect. For example, consider the effect of the outcome observed at trial t of a repeated choice task by an option generating payoff of +10 rarely (with probability 0.1) or -1 frequently (vs. the alternative of 0 with certainty). As exhibited in Fig. 1 (light blue[1] curves), if the outcome at t is +10 (rare good outcome), the option's choice rate (relative to its choice rate had the outcome been -1) initially increases, but very soon after (at $t + 3$), it decreases (to a point lower than that it started with), then gradually increases again, and only in the long term the effect diminishes.

Most attempts to develop descriptive models of learning focus on the long-term effects consistent with the common interpretation of the law of effect and ignore the short-term sequential effects. Indeed, most popular learning models assume recent outcomes have a relatively large effect on behavior, or positive recency effects (Busemeyer & Myung, 1992; Bush & Mosteller, 1955; Camerer & Ho, 1999; Denrell, 2005; Erev & Roth, 1998; Fudenberg & Levine, 1998; March, 1996; Sutton & Barto, 1998; Yechiam & Busemeyer, 2005). We believe that the main reason for the tendency to ignore the observed deviations from positive recency is the fact that the set of situations under which these deviations were documented is relatively narrow. All the clear deviations we are familiar with were observed in experimental studies that focus on simple tasks in which the environment is static and the decision makers receive full feedback. Specifically, in these studies, the feedback following each choice implied knowledge of the outcomes of both the selected option and the unselected option. Thus, decision makers had little reason to explore the environment. If deviations from positive recency emerge only when the feedback is complete but not under partial feedback conditions, which are more typical to the real world, it is safe to assume that these deviations reflect situation specific tendencies, rather than an intrinsic property of the learning process (Estes, 1962, 1964). Under one abstraction of this idea, the basic learning process is perfectly consistent with positive recency, but in certain settings (e.g., in simple tasks with complete feedback) people consider strategies that maximize return under the belief that the environment in about to change (Erev & Roth, 1999).

The main goal of the current research is to improve our understanding of the significance of the deviations from positive recency. The first part of the current analysis reanalyzes a wide set of published data that examine decisions from experience with limited feedback. In the second part of our analysis, we try to develop a simple model that can capture the observed short-term learning phenomena while allowing useful predictions of the long-term effects of experience.

1.1. The wavy recency effect and contingent average rules

The wavy recency effect of rare events, demonstrated in Fig. 1, is explained as the result of two distinct mechanisms (Plonsky et al., 2015). The first mechanism, which generates the initial short-term positive effect, captures the assumption that decision makers tend to respond to local trends in the outcomes. Specifically, a recent increase or decrease in observed payoffs produces a respective increase or decrease in the short-term attractiveness of the generating alternative. Plonsky et al. also show that this mechanism can lead to results previously attributed to the gambler's fallacy (the belief a change in outcomes "is due").

The second mechanism, which produces the wavy pattern, implies the use of "contingent average" (CA) rules. This mechanism can be a product of an adaptive reaction to the belief that past observed sequences tend to reappear. Specifically, Plonsky et al. (2015) suggest that decision makers first recall all previous experiences in the same contingency, experiences that followed the same sequence as the most recent sequence of outcomes, and then, they choose the alternative with the higher average payoff in these specific past experiences (i.e. with the higher contingent average). Consequently, following a rare outcome, the number of past experiences decision makers recall is small (there are only few past experiences which follow a sequence that includes a rare outcome). These few experiences, in turn, are not likely to include an experience that generated the rare outcome (i.e. reliance on small samples leads to underweighting of rare events, see e.g. Erev & Barron, 2005; Hertwig, Barron, Weber, & Erev, 2004; Hertwig & Erev, 2009; Rakow & Newell, 2010), and therefore following a rare outcome decision makers behave as if it becomes less likely that rare outcomes would be accounted for.

The use of CA rules has several attractive features (see Plonsky et al., 2015). First, in dynamic settings in which payoffs depend on the state of nature that changes in time, using them is highly effective. For example, when states of nature change according to a Markov chain, CA rules approximate the optimal policy. Second, CA rules provide a descriptive framework for binary repeated choice tasks capturing both the long-term aggregate choice rates and the short-term sequential choice dependencies. Finally, CA rules link choice behavior with a well-documented sensitivity to sequential patterns found in settings that range from perceptual-motor tasks to strategic 2-person games (e.g., Gaissmaier & Schooler, 2008; Nissen & Bullemer, 1987; Spiliopoulos, 2013). Such sensitivity has been shown to be robust across the human life-span

[1] For interpretation of color in Fig. 1, the reader is referred to the web version of this article.

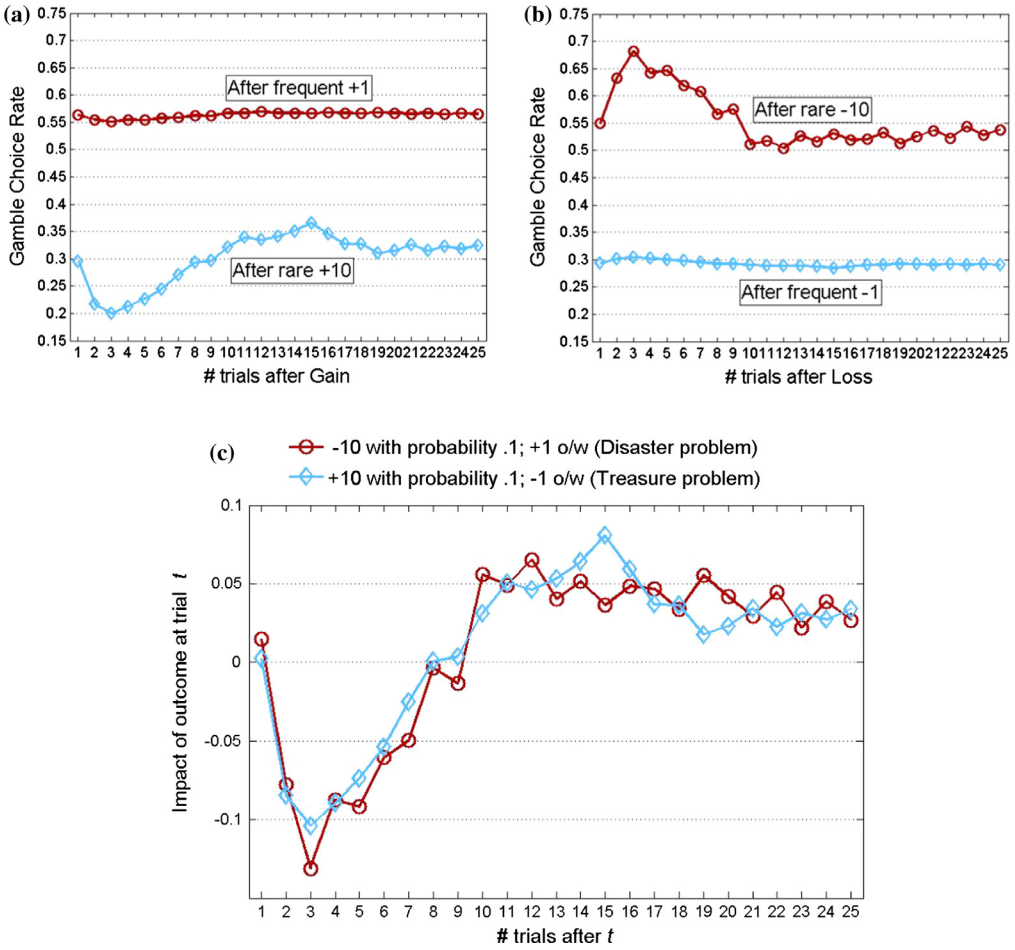


Fig. 1. The wavy recency effect of rare events in settings with complete feedback. Participants selected repeatedly for 100 trials between two unmarked buttons and received feedback concerning the payoff from both the chosen and the forgone option following each trial (see Appendix A). One option generated a payoff of 0 with certainty while the other was a risky gamble detailed in the legend. (a) Exhibits the choice rates of the gamble contingent on the gamble providing a gain at trial *t*; (b) exhibits the choice rates of the gamble contingent on the gamble providing a loss at trial *t*; and (c) presents the difference between the corresponding plots in (a) and (b). Thus the wavy curves in (c) reflect the impact of an outcome generated by the gamble at trial *t* on its choice rate in subsequent trials. Positive values imply "positive recency" and negative values imply "negative recency". Data is averaged across 48 participants from Nevo and Erev (2012) and 80 participants from Teodorescu et al. (2013), and originally exhibited in Plonsky et al. (2015).

(Weiermann & Meier, 2012), possibly already present as early as at the age of 3 months (Basirat, Dehaene, & Dehaene-Lambertz, 2014). Note in addition that the known tendency to expect outcome alternations to persist is also implied by the use of CA rules based on sequences of outcomes.

1.2. The current paper: three competing hypotheses

The current paper asks if the wavy recency effect that violates the positive recency assumption and signals an effort to discover patterns can be observed when the feedback is limited to the outcomes of the chosen alternative. Three competing hypotheses are examined here. The first hypothesis is that the lack of forgone outcome information leads decision makers to neglect the search for sequential patterns and focus on cognitively simpler strategies like those implied by leading models of learning. If this hypothesis is true, then the impact of rare outcomes on subsequent choice should be similar to that of non-rare outcomes. Specifically, we should find no evidence for a wavy recency effect. This hypothesis rests on the observation that it is much more difficult to discover patterns when the feedback is limited to the obtained outcomes.

When forgone feedback is provided, the information the decision maker acquires regarding the outcomes is independent of the choices he or she makes. Thus, detection of sequential patterns is possible even vicariously, when the decision maker chooses according to a very different strategy but suddenly notices that a pattern may exist. In contrast, when feedback is limited to the obtained outcomes, each choice made affects the information the decision maker has in the future (i.e. decisions entail an exploitation-exploration tradeoff). In this case, vicarious pattern spotting is much less likely. Indeed, in certain tasks, such as those we are dealing with in the current paper (see Appendix A, and Fig. 1), vicarious pattern spotting may be

even impossible. In such abstract tasks, it is reasonable for a decision maker to assume that in every trial, regardless of the choice made, outcomes are being generated by all possible options (even though the decision maker may only observe the outcome generated by the chosen option). This then means that if outcomes generated by a certain option follow a sequence or a trend, the decision maker would not be able to observe the pattern without choosing that option consecutively. Note this is in contrast to tasks in which the framing leads participants to believe the next outcome generated by each option "waits" until the decision maker selects the option – for example tasks in which options are framed as decks of cards (and the top card, with the outcome it implies, awaits until it is drawn). Moreover, without forgone feedback, the very definition of a sequence becomes vague. A sequence can be defined as the stream of recent outcomes observed, regardless of the option chosen, but such definition then implies that the very choices the decision maker takes change the sequential pattern, and it is not clear why this should be the case. Alternatively, a sequence may only be the stream of recent outcomes observed from a particular option, but then it is unclear how to treat the missing values that result from not choosing that option. Are they just ignored? Are they imputed with assumed values, and if so, in what manner?

A second competing hypothesis examined here rests on the assumption that the search for patterns is at the heart of human learning processes. Thus, with respect to the effect of outcomes on subsequent choice, feedback type does not matter: similar wavy recency effects should then be observed in settings with and settings without forgone feedback information. The third hypothesis is a refinement of the second hypothesis. It too implies that a search for patterns is a core characteristic of human learning, and as such emerges in settings both with and without forgone feedback. However, unlike the second hypothesis, it states that the lack of forgone feedback information influences short-term choice patterns in additional manners. Accordingly, the impact of rare outcomes in settings without forgone feedback would, on one hand, reflect a use of CA rules, but on the other hand, differ from the impact of rare outcomes in settings with forgone feedback. One robust phenomenon that emerges in settings with partial information and that has a potential to alter the observed impact of rare events is the "hot stove effect" (Denrell, 2007; Denrell & March, 2001; and see Einhorn & Hogarth, 1978). This effect is named after a story by Mark Twain, telling of a cat who once sat on a hot stove and never again sat on the stove, whether it was hot or cold. More specifically, this effect implies that bad outcomes in settings without forgone feedback would likely have a stronger and more long-lasting effect on subsequent choice than in settings with forgone feedback.

To compare these three hypotheses, we analyze data from four published studies (Camilleri & Newell, 2011; Di Guida, Erev, & Marchiori, 2015; Erev et al., 2010; Nevo & Erev, 2012) that used the paradigm presented in Appendix A (and see Fig. 1). First, we focus on 16 problems ran with forgone feedback and examine the effect of only the obtained outcomes on subsequent choice (rather the effect of either the obtained or the forgone outcomes). The results of this analysis reveal a pattern that we previously overlooked, when aggregating the effects of both obtained and forgone outcomes: The wavy recency effect of rare positive events is different from the wavy recency effect of rare negative events.

Second, we compare these effects to the effects of outcomes on choice in 16 problems identical to the first 16, with the exception that they were run without forgone feedback. The results show similar, but not identical, short-term effects, in support of our third hypothesis above. Finally, we develop a model of choice in settings without forgone feedback, an extension of the model proposed by Plonsky et al. (2015). The model is fitted based on the long-term aggregate choice rates of 60 problems (the Estimation Set from Erev et al., 2010), and evaluated based on its predictive value both for the long-term aggregate choice rates of 60 different problems (the Competition Set from Erev et al., 2010) and for the short-term impact of observed outcomes on subsequent choice. The results show that this model outperforms the models proposed in previous attempts to capture this large data set, and captures reasonably well the individual heterogeneity in choice. Moreover, we show a small variant of the proposed model can achieve all that, as well as capture behavior in a different task in which outcomes are auto-correlated.

2. The effect of rare obtained outcomes on subsequent choices

Previous demonstrations of the wavy recency effect of rare events, which were restricted to problems with forgone feedback, did not distinguish between the effects of obtained and the effects of forgone outcomes (see e.g. Fig. 1). As we focus in this study on the effect of rare events in settings that do not include forgone feedback, it is important to first examine whether the effect of obtained rare outcomes differs from the effect of forgone rare outcomes in settings with complete feedback. We can then compare the effect of rare obtained outcomes in these complete feedback settings to their effect in settings with incomplete feedback.

2.1. Surprise-triggers-change

One potentially important phenomenon that may generate a different effect of obtained and forgone outcomes is "sur prise-triggers-change" (Nevo & Erev, 2012; and see Newell, Rakow, Yechiam, & Sambur, 2015). According to this short-term phenomenon, agents are much more likely to switch their choice (from that made in the previous trial) after observing a rare (and surprising) outcome. This effect is consistent with neural research suggesting increased activation with prediction error (which is likely to be larger when observing surprising outcomes; Schultz, 1998). Nevo and Erev show that the increased tendency to switch after surprising outcomes is more predictive of choice than the assumption that people tend to select the alternative providing the best outcome in recent trials. That is, it is more predictive than the assumption of

Table 1
Problems analyzed in Section 2.

Type	Safe	Risky			With forgone		Without forgone	
		High	Prob. High	Low	*n*	Source	*n*	Source
Treasure	7	16.5	0.01	6.9	27	NE12	20	EEA10
Treasure	−9.4	−2	0.05	−10.4	29	NE12	20	EEA10
Treasure	−4.1	1.3	0.05	−4.3	29	NE12	20	EEA10
Treasure	−18.7	−7.1	0.07	−19.6	29	NE12	20	EEA10
Treasure	−25.4	−8.9	0.08	−26.3	29	NE12	20	EEA10
Treasure	−7.9	5	0.08	−9.1	29	NE12	20	EEA10
Treasure	0	10	0.1	−1	31	DiG15	20	DiG15
Treasure	11.5	25.7	0.1	8.1	29	NE12	20	EEA10
Treasure	2	14	0.15	0	20	CN11	20	CN11
Disaster	−3	0	0.9	−32	20	CN11	20	CN11
Disaster	0	1	0.9	−10	31	DiG15	20	DiG15
Disaster	9	10	0.9	0	20	CN11	20	CN11
Disaster	2.2	3	0.93	−7.2	29	NE12	20	EEA10
Disaster	25.2	26.5	0.94	8.3	29	NE12	20	EEA10
Disaster	6.8	7.3	0.96	−8.5	28	NE12	20	EEA10
Disaster	11	11.4	0.97	1.9	29	NE12	20	EEA10

Note. Each problem was a choice between a Safe option (a certain payoff) and a Risky option that provided High with probability Prob. High and Low otherwise. Analyzed data is adopted from Camilleri and Newell (2011) (CN11); Di Guida et al. (2015) (DiG15); Erev et al. (2010) (EEA10); and Nevo and Erev (2012) (NE12).

positive recency. For example, when repeatedly selecting between the status quo (payoff of zero with certainty) and a bad gamble providing −10 with probability 0.1 and +1 otherwise (denoted "(−10, 0.1; +1)"; see Fig. 1), a surprising −10 in trial *t* decreases choice rates of the gamble for agents who selected the gamble in trial *t*, but increases the choice rate of the gamble for agents who selected the status quo in trial *t*. Similarly, observing a surprising +10 when choosing between the status quo and (+10, 0.1; −1), increases the risky choice rate of the gamble for agents who only just selected status quo but decreases it for agents who only just selected the gamble (and received a payoff of +10).

These two examples illustrate that regardless of their value, rare obtained outcomes tend to reduce the choice rate of the chosen alternative that generates them. Consequently, the impact of rare and bad obtained outcomes ("disasters") is in line with many learning models' assumption of positive recency (a bad outcome decreases the choice rate of that option). Yet, the impact of rare and good obtained outcomes ("treasures") is at odds with this assumption (a *good* outcome decreases the choice rate of that option). Therefore, henceforth, we analyze the effects of obtained outcomes on subsequent choice separately for treasure settings (rare good outcomes) and for disaster settings (rare bad outcomes).

2.2. Analysis of published data

2.2.1. Method

In this section, we analyze data adopted from four published studies. Two of the studies (Camilleri & Newell, 2011; Di Guida et al., 2015) include both problems with and problems without forgone feedback; Nevo and Erev (2012) include only problems with forgone feedback; and Erev et al. (2010) include only problems without forgone feedback. All four studies used variants of the clicking paradigm presented in Appendix A. Additional details regarding the method used in each of these studies are provided in the respective papers. Hereinafter we briefly review the essentials.

2.2.1.1. Problem selection. The studies the data is taken from include many choice tasks. To allow for a fair and effective comparison between the effects of rare outcomes in settings with and settings without forgone feedback, we chose to focus only on problems that include rare events (defined here as occurring with probability not greater than 0.15) and that were run in both settings. Sixteen pairs of problems were found to hold these criteria (each pair includes the same problem run, independently, in both types of settings; see Table 1).

2.2.1.2. Participants. Between 20 and 31 students faced each problem (see Table 1 for details). Each participant faced either only problems with or only problems without forgone feedback. In each study, each participant faced all the problems of the same type taken from that study.[2] In total, data of 80 participants in the full information condition and 60 participants in the partial information condition is analyzed.

2.2.1.3. Materials and procedure. In all cases, participants were instructed to repeatedly select between two unmarked alternatives and were not told anything regarding the payoff distributions. In practice, all problems were a repeated choice, for at

[2] In Nevo and Erev's (2012) study, the data for 2 subjects in one problem and another subject in a different problem could not be recovered, hence the somewhat different sample sizes in two of the problems.

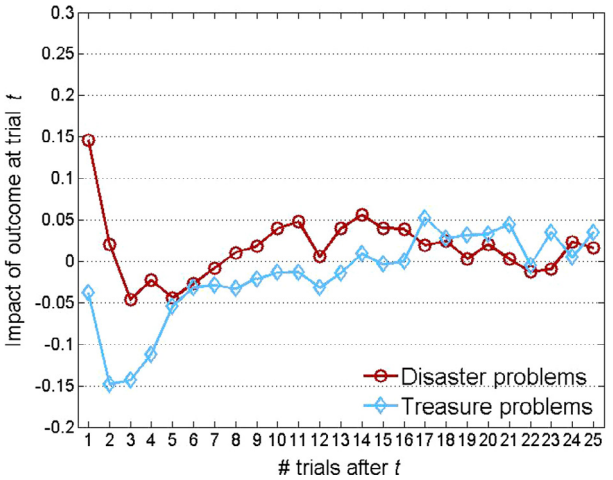


Fig. 2. Average impact of an obtained outcome at trial t on subsequent choices in problems that include forgone outcomes feedback. The impact at trial $t + n$ is the difference between the choice rate of the Risky option contingent on obtaining a high payoff at trial t and its choice rate contingent on obtaining a low payoff at trial t. The dark red curve is the average of 7 disaster (i.e. rare low payoff) problem curves, and the light blue curve is the average of 9 treasure (i.e. rare high payoff) problem curves (see Table 1). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

least 100 trials, between a certain amount (a safe alternative) and a binary gamble that included a rare event. In problems which included more than 100 trials, only the first 100 trials are analyzed. Each choice was immediately followed by the presentation of the outcome realized by the chosen alternative in that trial. In problems with complete information, each choice was additionally followed by the presentation of the outcome realized by the forgone alternative (see Appendix A). All decisions made were consequential for the participants' final monetary payoff.

2.2.2. Results

2.2.2.1. Problems with complete feedback information. We computed, separately for treasure problems and for disaster problems, the average impact of an outcome obtained from the Risky alternative at trial t on the choice rate of the Risky alternative in trials $t + 1$ through $t + 25$ (Fig. 2). In particular, the impact here is only computed for trials at which the subjects chose the Risky alternative and for trials 25–100 in each problem. The impact at trial $t + n$ is defined as the difference between the choice rate of the Risky alternative contingent on a high (good) outcome at trial t and the choice rate of the Risky alternative contingent on a low (bad) outcome at trial t.[3] For example, consider a participant facing 100 repeated choices between zero with certainty and the gamble (+10, 0.1; −1). To compute the impact of an outcome at trial $t + 1$ on that participant's subsequent choice, we compute (a) the proportion of risky choices across all trials number 25–100 that immediately followed a + 10 outcome and (b) the proportion of risky choices across all trials number 25–100 that immediately followed a −1 outcome. The difference between these two proportions is the impact at $t + 1$. The impact at trial $t + 2$ is the difference between the risk rate in trials number 25–100 that occur exactly two trials after observing +10, and the risk rate in trials number 25–100 that occur exactly two trials after observing −1, and so on. Hence, a positive impact score at trial $t + n$ implies a positive recency effect of the outcome at trial t, whereas a negative impact score at trial $t + n$ implies a negative recency effect (which means better outcomes reduce the attractiveness of the gamble compared to worse outcomes).

Fig. 2 illustrates the surprise-triggers-change phenomenon. The initial impact (at trial $t + 1$) clearly differs between treasure problems and disaster problems. In disaster problems, it is highly positive, which implies that immediately after obtaining a rare bad outcome agents tend to avoid the gamble. Conversely, in treasure problems, the initial impact is negative, which implies agents also tend to avoid the gamble immediately after getting a rare good outcome. To statistically test this difference, we created, for each individual subject, two impact curves, one for each type of problem (treasures and disasters). This is done by aggregating the contingencies from all problems of the same type faced by that individual.[4] We then

[3] Because relatively few participants faced each problem, because we are dealing with rare events, and because we are focusing on only some of the trials in each problem (those at which participants selected the Risky option), there are generally few observations per participant per problem. To face with the relatively high noise involved, the presented impact curves in this study are the average aggregated impact curves. Specifically, we first compute each participant's impact curve in each problem he or she faced, then average all individual curves in each problem to generate an aggregate curve for that problem (ignoring missing values, e.g., for an individual who did not observe both possible outcomes from the Risky option), and finally average over all aggregate curves in each type of problem to create the presented curve. The online supplemental material shows that impact curves averaged across individual participants are noisier but qualitatively similar.

[4] This aggregation is meant to reduce the number of missing data points in the impact curve. A missing value for a certain participant in a certain problem implies, e.g., that the participant did not observe both possible outcomes from the Risky option in that problem. Even with this aggregation method, not every participant has two full (all 50 contingencies) data point sets, although the majority do. Thus, the individual-level analyses henceforth represent only those participants for which all relevant impact curve points could be computed. Specifically, for the current analysis regarding the impact at $t + 1$, we use all participants that have available an impact score at $t + 1$ for both treasure and disaster problems.

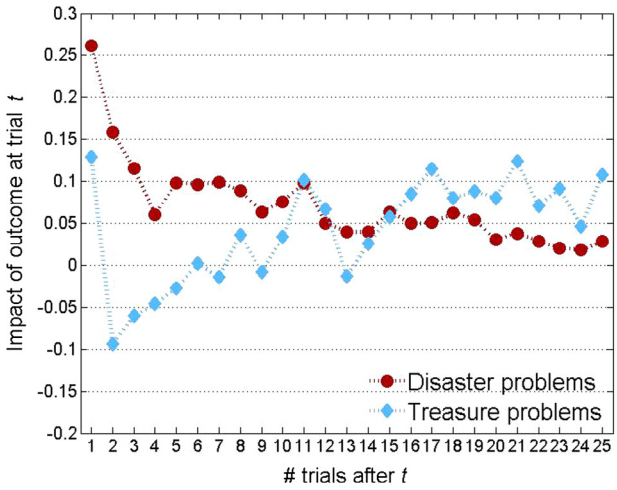


Fig. 3. Average impact of an obtained outcome at trial t on subsequent choices in problems that include feedback only for the obtained outcomes (and not for forgone outcomes). The impact at trial $t + n$ is the difference between the choice rate of the Risky option contingent on obtaining a high payoff at trial t and its choice rate contingent on obtaining a low payoff at trial t. The dark red curve is the average of 7 disaster (i.e. rare low payoff) problem curves, and the light blue curve is the average of 9 treasure (i.e. rare high payoff) problem curves (see Table 1). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

computed the difference between the initial impact of an outcome in disaster problems and this impact in treasure problems. The mean difference of 0.208 ($SD = 0.36$) is highly significant, $t(55) = 4.27$, $p < 0.001$. Moreover, 71% of participants for whom this difference can be computed (40 of 56; sign test $p = 0.001$) exhibit a higher impact at $t + 1$ in disaster problems than in treasure problems.

Note that—assuming the observed difference is due to surprise-triggers-change—had we aggregated the impact curves following both obtained and forgone outcomes (as in Plonsky et al., 2015, and Fig. 1), there should have been no difference between the curves of treasure problems and disaster problems. Accordingly, Fig. S.1 (see the online supplementary material) shows the mean impact curves for treasures and disasters are nearly identical when considering both obtained and forgone outcomes.

Additionally, Fig. 2 shows that for both treasure problems and disaster problems, outcomes have a wavy impact on subsequent choice. While the initial impact is relatively high, very soon after (within 2–3 trials) it drops to a trough, then increases gradually, and finally, in the long-term, the impact more or less stabilizes. This wavy pattern results from the choice rates that follow rare events and does not emerge in settings without them (not shown, but see Plonsky et al., 2015). To statistically test the wavy pattern, we compared the average impact of an outcome on choices 2–4 trials after it occurs with its average impact on choices 12–14 trials after it occurs.[5] In accordance with the wavy recency prediction, the impact 2–4 trials after the outcome is significantly lower than the impact 12–14 trials after the outcome both for treasure problems ($M = -0.144$, $SD = 0.26$, $t(55) = -4.16$, $p < 0.001$) and for disaster problems ($M = -0.051$, $SD = 0.22$, $t(75) = -2.06$, $p = 0.043$). Yet, while in treasure problems 68% (38 of 56, sign-test $p = 0.002$) of participants exhibit the effect, in disaster problems only 50% (38 of 76, sign-test $p = 0.64$) of participants exhibit it (33 of them exhibit a reversed pattern and 5 show no effect). The difference between the proportions is statistically significant, $z = 2.052$, $p = 0.040$. This suggests that when accounting for only the obtained outcomes, the wavy effect is more robust in treasure problems than in disaster problems.

2.2.2.2. Problems with incomplete feedback information. The aggregated impact curves of the 16 problems with incomplete feedback, for treasure and for disaster problems separately, are exhibited in Fig. 3. Similarly to problems with complete feedback, the immediate impact of an outcome in treasure problems differs from the immediate impact of an outcome in disaster problems. Testing this difference for the 33 participants for whom the immediate impact in these settings can be reliably computed (those who made at least one choice after each possible outcome from the risky choice in both disaster and treasure problems) reveals the mean difference of 0.178 ($SD = 0.36$) is significant, $t(32) = 2.81$, $p = 0.008$. Moreover, this difference is positive for 73% of these participants (sign-test $p = 0.014$). This observation suggests that surprising outcomes tend to increase the likelihood of change in settings with complete as well as settings with incomplete feedback.

Moreover, like in problems with complete feedback, outcomes in treasure problems exhibit a wavy recency pattern: After an initial large impact, the impact drops to a trough, then begins to slowly increase and only in the long-term stabilizes. Particularly, the average impact of the outcome at trial t on the choice at trials $t + 2$ through $t + 4$ is significantly lower than

[5] Plonsky et al. (2015, and see Fig. 1) show that the decision makers' sensitivity to a rare event is minimal three trials after it occurs and is maximal around 12–15 trials after it occurs, hence our choice of trials to average in order to test the wavy recency effect.

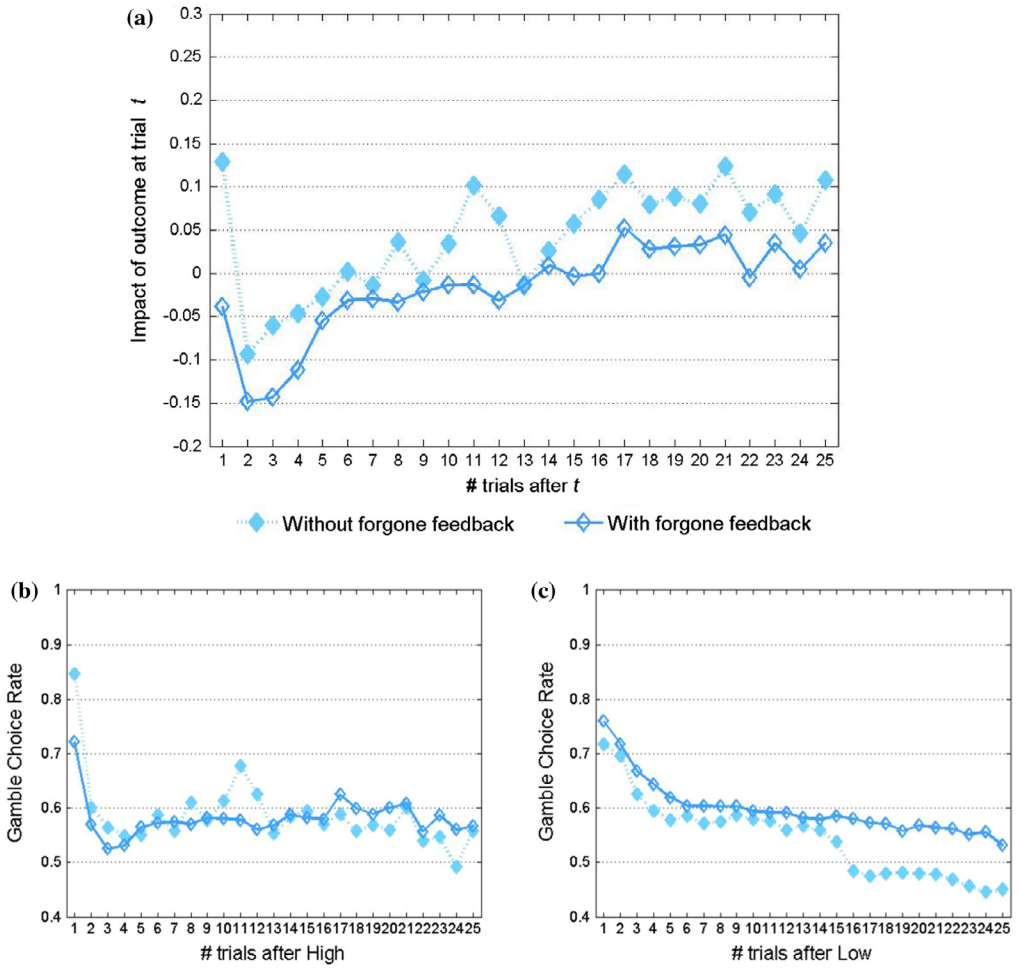


Fig. 4. The wavy recency effect in treasure problems, with and without forgone feedback. The impact of an obtained outcome at trial t on subsequent choices (a) is the difference between the choice rate of the risky gamble option contingent on a High payoff at trial t (b) and the choice rate of the risky gamble option contingent on a Low payoff at trial t (c). Plots are averages of the 9 treasure problems exhibited in Table 1.

its impact on the choice at trials $t + 12$ through $t + 14$ ($M = -0.129$, $SD = 0.33$, $t(34) = -2.33$, $p = 0.026$). Indeed, this pattern is evident for 77% of the participants for whom it is possible to compute (27 of 35, sign-test $p < 0.001$)

However, unlike problems that include forgone outcomes information, in disaster problems there does not seem to be any evidence for a wavy recency effect. The dark curve in Fig. 3 suggests an almost monotonic decrease in the impact of outcomes in such problems. This pattern, which is consistent with the common assumption of positive recency, suggests that the initial impact of an obtained outcome in disaster problems without forgone information is largest in the first trial after it occurs and the effect gradually diminishes in time. Indeed, the average impact 2–4 trials after the outcome is larger than the average impact 12–14 trials after the outcome ($M = 0.020$, $SD = 0.27$). This difference is insignificant, $t(56) = 0.587$, $p = 0.559$, but the fact it is positive implies the pattern is opposite than that predicted by the wavy recency account. In fact, in these disaster problems, this difference is positive for the majority of participants (34 of 57) though the two-sided sign-test is only marginally significant, $p = 0.076$.

2.2.2.3. Comparison of problems with complete and with incomplete feedback information. The results above suggest that the effect of rare obtained outcomes in settings with and settings without forgone feedback is similar in treasure problems but different in disaster problems. Fig. 4a (for treasure problems) and Fig. 5a (for disaster problems) affirm this statement by facilitating the comparison of the two setting types (with and without forgone feedback) separately for treasure and for disaster problems. In addition, they show that both for treasure and for disaster problems, the impact curves in the no-forgone settings are set above the impact curves in the with-forgone settings. This suggests that the impact of an obtained outcome is larger, and more long-lasting, when the provided information is limited to the obtained outcomes.

Comparison of Figs. 4b, 4c, 5b, and 5c reveals the reason for the higher impact in settings without forgone feedback. These figures present the choice rates of the risky alternatives contingent on either obtaining a relative gain (Figs. 4b and 5b) or

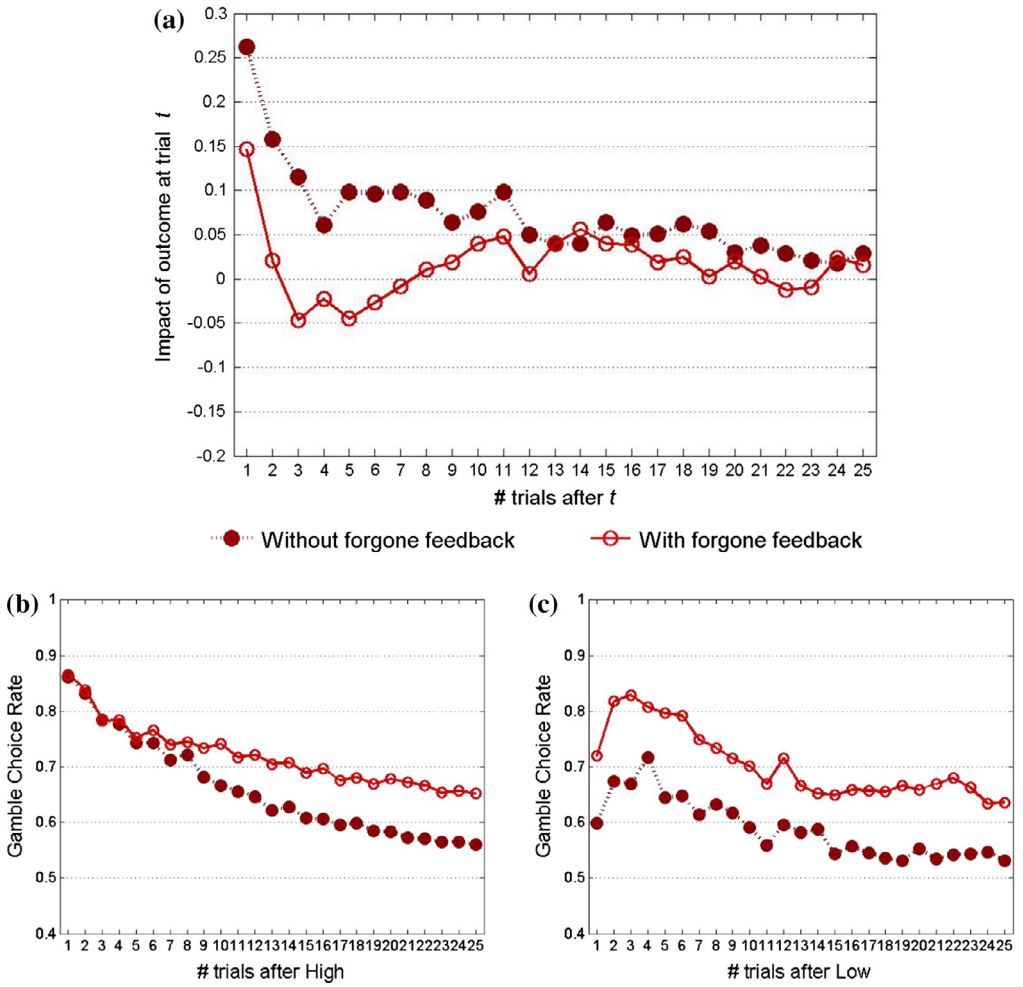


Fig. 5. The wavy recency effect in disaster problems, with and without forgone feedback. The impact of an obtained outcome at trial t on subsequent choices (a) is the difference between the choice rate of the risky gamble option contingent on a High payoff at trial t (b) and the choice rate of the risky gamble option contingent on a Low payoff at trial t (c). Plots are averages of the 7 disaster problems exhibited in Table 1.

obtaining a relative loss (Figs. 4c and 5c) at trial t. That is, they break down each of the impact curves (from Figs. 4a and 5a) to their two parts (e.g., the curves in Fig. 4a are the difference between the respective plots in Fig. 4b and c). For illustration, one may consider Fig. 4b as representative of the choice rate of the gamble (+10, 0.1; −1) over a safe 0 as a function of the number of trials since obtaining rare +10, whereas Fig. 4c represents the choice rate this gamble as a function of the number of trials since obtaining −1. Similarly, Fig. 5b represents the choice rate of the gamble (−10, 0.1; +1) over a safe 0 as a function of the number of trials since obtaining +1, whereas Fig. 5c represents the choice rate of this gamble as a function of the number of trials since obtaining rare −10.

As Figs. 4c and 5c show, bad outcomes (relative losses) lead to generally less choice of the Risky option when forgone feedback is not provided, while the choice rates contingent on gains (Figs. 4b and 5b) are considerably more similar. In other words, the asymmetric information creates a "hot stove" effect: the "bad impression" created by the loss at trial t reduces the tendency to select the risky option when forgone feedback is not provided and there is less chance that this impression would be rectified in subsequent trials. In fact, the hot stove effect can also hint for the reason the curves in Fig. 5b get further apart starting at some point: because losses are more likely to have appeared the further the curve is from its origin, the choice rates in the no forgone setting get increasingly more likely to diminish in time.

In addition, Figs. 4b and 5c substantiate the claim that the wavy pattern is due to the effect of rare events. Both in treasure problems (Fig. 4b) and in disaster problems (Fig. 5c), the choice rates contingent on the rare events are "wavy" (whether forgone feedback is provided or not). For example, maximal underweighting of the rare event is exhibited around 3 trials after it occurs. However, following a frequent event (Figs. 4c and 5b), the choice rates are much more monotonic. Finally, the fact that experiencing a frequent loss is associated with *more* risk seeking immediately after it is obtained, an effect which decays with time (Fig. 4c), suggests that inertia may play an important role in these settings, as has been previously suggested (Biele, Erev, & Ert, 2009).

To statistically test the differences between the effect of outcomes on subsequent choice in settings with and settings without forgone feedback, we implemented (using SAS 9.4 procedure *MIXED*) two linear mixed models—one for treasure problems and one for disaster problems—with the impact of an outcome at trial t on subsequent trials (for trials $t+1$ through $t+25$) as the dependent variable. The predictors used in each model were setting type (either with or without forgone feedback), number of trials since the occurrence of the outcome (between 1 and 25) and the interaction between them. In addition, because the initial immediate impact (i.e. at $t+1$) of an outcome is clearly different than its impact in subsequent trials (e.g., see Fig. 2), each model also included a dummy for the initial impact and for its interaction with the setting type. We used correlated error terms for the 25 different observations of the same subject, and specifically, completely unstructured covariance matrices.[6] Degrees of freedom were obtained using the Satterthwaite approximation. Restricted maximum likelihood estimations are reported.

The analyses show both models are highly significant, $\chi^2(324) = 1401$, $p < 0.001$ and $\chi^2(324) = 2437.3$, $p < 0.001$, for treasure and for disaster problems respectively, and their results corroborate the observations stated above. For the treasure model (fitted for the data presented in Fig. 4), the main effect for the number of trials since the outcome is significant, $F(1,87) = 10.27$, $p = 0.002$, but its interaction with setting type is not, $F(1,87) = 1.20$, $p = 0.28$, indicating that the development of impact over time is similar in both setting types. The same is true for the effect of the initial impact which is significant $F(1,87) = 14.83$, $p < 0.001$, while its interaction with setting type is not $F(1,87) = 0.62$, $p = 0.43$. Moreover, a significant main effect of setting type, $F(1,87) = 9.76$, $p = 0.002$, implies a hot stove effect.

For the disaster model (fitted for data presented in Fig. 5), in contrast, the interaction between the number of trials since the outcome and the setting type is significant: $F(1,129) = 4.60$, $p = 0.034$. This suggests that in these problems the development of impact over time when forgone feedback is provided is different than its development when forgone feedback is not provided. This difference is not due to the initial impact: while its main effect is significant, $F(1,129) = 55.76$, $p < 0.001$, its interaction with setting type is not, $F(1,129) = 0.33$, $p = 0.57$. Finally, the main effect for setting type is significant in these problems as well, $F(1,129) = 8.73$, $p = 0.004$, which implies a hot stove effect. The full fixed effects tables can be found in the online supplemental material.

3. Robustness of the wavy pattern and a descriptive model

The analyses presented in Section 2 focused on 16 problems that were studied in settings with and settings without providing the subjects information concerning the forgone payoffs. The results show that wavy recency effect of rare events emerges even when the feedback is limited to the obtained payoffs, at least in treasure problems. In addition, the observed response appears to reflect an interaction of this effect with two additional short-term learning phenomena: surprise-triggers-change, and the hot stove effect. The current section takes two related steps toward clarifying the implications of this interaction in the context of learning with limited feedback. The first is an examination of the robustness of the interaction by studying a wider data set. The second step is an evaluation of the relationship between the current short-term phenomena, and the long-term phenomena documented in previous studies of decisions from experience. In order to achieve this goal, we examine whether a model that abstracts the interaction between the three short-term effects can reproduce, with the same set of parameters, the aggregate choice rates (long-term results) in 120 randomly selected choice problems.

3.1. The experiment

The data analyzed in this section is taken from the Technion Prediction Tournament (TPT; Erev et al., 2010; see Appendix B), a modeling competition in which the organizers, who ran the experiments, made publicly available data of 60 choice problems (*Estimation set*) and challenged other researchers to develop a model predicting the results of a *Competition set*.[7] The Competition set included 60 different problems that are members of the same problem space from which the problems in the Estimation set were derived. Specifically, all 120 problems were repeated choice tasks without forgone feedback. The winner of the TPT was the model that provided the lowest (best) mean squared deviation (MSD) from the 60 aggregate choice rates of the Competition set.

3.1.1. Method

Each of the 120 problems was a 100-trial repeated choice between two unmarked alternatives. As in Section 2, one alternative was safe and provided a medium payoff with certainty, while the other alternative was risky and provided either a

[6] This was done because each point in the individual-subject impact curves represents aggregated contingencies, and thus the inter-class correlations within each curve may be of any kind (and in particular, negative). We also tested for other covariance matrix structures (compound symmetry, autoregressive, heterogeneous autoregressive and Toeplitz) and log-likelihood tests, as well as Akaike Information Criterion, have shown the unstructured matrix is best. Finally, as previously noted, we only used subjects whose impact curves did not contain any missing data points. However, we also tested the models including any existing data points of subjects with partial curves and the qualitative results were practically identical.

[7] The 11 problems from Erev et al.'s (2010) study, analyzed above, belong to the estimation data set. These problems are used in this section too. To reduce the risk of overfitting these problems we evaluate the model (Section 3.2) based on its ability to predict the results in the competition data set (which these problems do not belong to). We also verified that the results presented in Section 3.1.2 hold when these 11 problems are excluded.

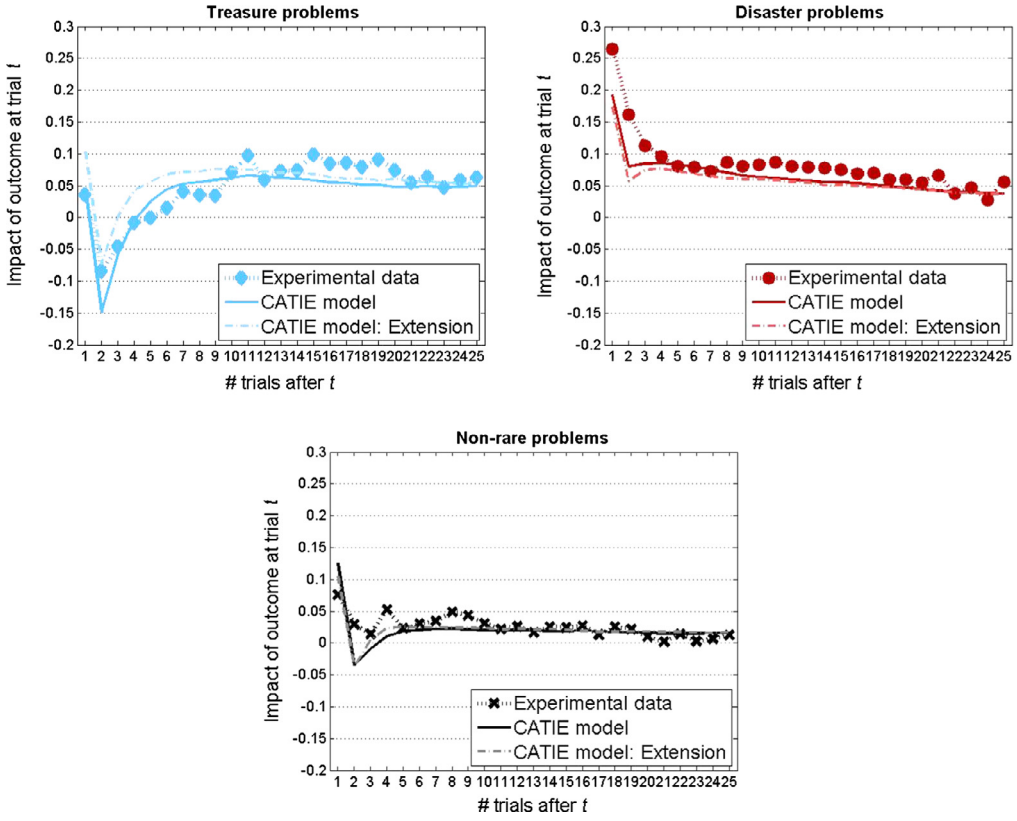


Fig. 6. Experimental and predicted impact of an outcome on subsequent trials in problems without forgone feedback. The impact at trial $t+n$ is the difference between the choice rate of the Risky option contingent on obtaining a High payoff at trial t and its choice rate contingent on obtaining a Low payoff at trial t. Data is adopted from the Technion Prediction Tournament (Erev et al., 2010, and see Appendix B). The top left panel exhibits the average impact curves for 43 treasure problems (problems with rare High payoff); The top right panel exhibits the average impact curves for 48 disaster problems (problems with rare Low payoff); and the bottom panel exhibits the average impact curves for 29 problems without rare events. CATIE is a contingent average, trend, inertia, and exploration model with the parameters $\tau = 0.29$, $\varepsilon = 0.3$, $\varphi = 0.71$, $K = 2$. The parameters for the extension of CATIE (Section 3.2.2.1) are $\tau = 0.30$, $\varepsilon = 0.48$, $\varphi = 0.78$, $K = 2$, $\delta = 0.55$.

high or a low outcome. A predefined algorithm randomly selected the exact parameters of each problem (i.e. the medium, low, and high payoffs, and the probability for a high payoff). The data set therefore represents a wide range of problems.

Twenty participants faced 12 problems each. Thus, in total data from 200 participants is analyzed. Following each choice, participants received as feedback the realized outcome from the chosen alternative (but not from the forgone alternative). At the end of the experiment, participants received the obtained payoff of one randomly selected trial plus a show up fee. The 120 problems and the aggregate choice rates are presented in Appendix B. Further details regarding the method used in the experiment is provided elsewhere (Erev et al., 2010).

3.1.2. Results

The aggregate choice rates for each problem are shown in Appendix B. Following the analysis above that shows treasure and disaster problems lead to very different impact curves, we grouped the 120 problems into three groups: treasure problems, disaster problems, and problems that are neither treasure nor disaster problems (i.e. problems that do not include rare events). To compute the average impact curve in each group, we first computed the mean impact curve for each participant in each problem, then created an average impact curve for each problem (ignoring missing values) and finally averaged all these problems impact curves in each group. The dotted plots in Fig. 6 depict the resulting curves.

The results suggest that the observations made in Section 2 are robust. The average impact of an outcome in a treasure problem on subsequent choice is initially slightly positive, then becomes sharply negative, then gradually increases to reach a point higher than that it started with, and in the long-term the effect diminishes. In disaster problems, the immediate impact of an outcome on the choice in the next trial is far greater than in treasure problems, corroborating the evidence for a surprise-triggers-change phenomenon. Moreover, the impact in disaster problems gradually decays in time and does not exhibit a wavy pattern. Thus, the difference between the impact of rare events in treasure and in disaster problems from Section 2 replicates in the TPT data as well.

The average impact of an outcome in a problem without rare events differs from that of outcomes in problems with rare events.[8] Like the impact in disaster problems (and unlike that in treasure problems), the maximal impact of the outcome in trial t in non-rare problems is obtained in trial $t + 1$. Yet, unlike the impact in disaster problems (and like that in treasure problems), the impact does not then decreases monotonically. Instead, it drops to a small trough, then slightly increases and the effect fades in time.

3.2. Contingent average, trend, inertia, and exploration (CATIE) model

Assuming that the three short-term choice phenomena demonstrated above reflect central properties of human learning, a model that captures these phenomena should allow the prediction of the main long-term learning phenomena. In order to evaluate this optimistic hypothesis, we chose to develop a model that refines two existing models: *CAT* (Contingent Average and Trend), which was developed to capture behavior in binary choice problems with forgone information (Plonsky et al., 2015), and *4-Mode Contingent Sampler* (Biele et al., 2009), which was developed to capture behavior in dynamic settings with high payoff autocorrelation (a model simplifying the Instance Based Learning model; Gonzalez, Lerch, & Lebiere, 2003).

The current model, referred to as "Contingent Average, Trend, Inertia, and Exploration" (CATIE), assumes that in the first two trials of the experiment the subject chooses each of the two alternatives. Then, starting at Trial 3, CATIE, like the 4-mode contingent sampler, chooses according to one of four modes: exploration, exploitation, inertia, and a use of a simple heuristic.

As in CAT, the model assumes that when people are in heuristic mode, which happens with probability τ (a free parameter), they tend to follow trends in the observed payoffs. In our no-forgone settings, a trend exists if the participant just selected the same alternative for the second consecutive time and the observed payoffs in these two times were different (see Hohle and Teigen (2015), for evidence that two periods are sufficient to define a trend). The heuristic implies that people would prefer to choose an option after it displays an increasing trend but avoid it after it displays a decreasing trend. The trend is increasing if the last payoff is higher than the payoff preceding it and is decreasing otherwise.[9]

If the model has not made a decision using the heuristic mode (i.e. with probability $1 - \tau$ if a trend is observed and always if a trend is not observed), it assumes the agent may enter an exploration mode which implies random choice. The probability of entering the exploration mode is a function of the surprise the agent experiences: observing surprising outcomes leads to increased exploration. Specifically, probability of exploration in trial $t + 1$ (contingent on not using the heuristic mode) equals:

$$p(Explore|No\_Heuristic)_{t+1} = \varepsilon \cdot \frac{1 + Surprise_t + Mean\_Surprise}{3} \tag{1}$$

where $\varepsilon \in [0, 1]$ is a free parameter capturing a basic exploration tendency, $Surprise_t \in [0, 1]$ is the surprise experienced in trial t and *Mean_Surprise* is the average of experienced surprise in trials 1 through t. Thus, the probability of exploration is minimal (equals $\varepsilon/3$) when the surprise is minimal (both $Surprise_t$ and *Mean_Surprise* equal 0), and maximal (approaches ε) when the surprise is large.

The surprise in trial t is computed as:

$$Surprise_t = \begin{cases} \frac{|Exp_t^i - V_t^i|}{ObsSD^i + |Exp_t^i - V_t^i|}, & \text{if } ObsSD^i > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where Exp_t^i is the payoff the model expects to observe after choosing alternative i in trial t, V_t^i is the payoff the model actually observes after choosing alternative i in trial t, and $ObsSD^i$ is the standard deviation of all payoffs observed from alternative i in trials 1 through t. This function captures the intuition that surprise increases with the difference between the expected and the realized outcomes (Nevo & Erev, 2012; Rescorla & Wagner, 1972; Teigen & Keren, 2003) but decreases when the subject expects to see significant variability in these outcomes (as a result of experiencing a lot of variability in the past). The payoff expectation Exp_t^i for alternative i is its contingent average (the average payoff obtained in the past given the current contingency; see below). The addition of the *Mean_Surprise* term to the probability of exploration captures the intuition that although safe alternatives are never surprising, exploration is expected to be relatively high even after choosing the safe alternative in case the overall setting has been surprising in the past.

In case the model has not made a decision using either the heuristic or the exploration modes, it enters an inertia mode with probability φ (a free parameter). This mode prescribes repetition of the previous trial choice (Biele et al., 2009; Gonzalez & Dutt, 2011).

Finally, if the model has not made a choice in one the former three modes, it enters the exploitation mode in which it uses a CA rule based on k recent outcomes (CAB-k; Plonsky et al., 2015). That is, it chooses the option with the higher contingent average defined as the average payoff observed in all previous trials which followed the same sequence of k outcomes as the

[8] In the datasets we use (and naturally, in most datasets), rare events are also extreme outcome events. Thus, we cannot rule out the option that the difference between the observed impact curves in problems with and problems without rare events is due to the extremity of the outcomes, and not a result of the rarity of the outcomes (see Ludvig, Madan, and Spetch (2014) for a discussion on extreme events in experiential choice). Yet, note that the abstraction of the model we present in the Section 3.2 (and that captures well the results) assumes the effect is due to the probability, rather than the extreme magnitude of outcomes.

[9] Notice this heuristic does not imply best response to the last outcome (a win-stay-lost-shift strategy). For example, it increases choice following a high outcome when it follows a low outcome, but not when it follows another high outcome (Hohle & Teigen, 2015).

most recent sequence.[10] The model assumes the value of k is selected from a uniform distribution of integers between zero and K (a free parameter). In particular, $k = 0$ implies all previous experiences are averaged. If there are no past experiences which followed the same sequence length k (a new contingency), the model assumes the current contingency is most similar to another already observed contingency (randomly selected) and assigns its contingent average to the current trial.[11]

CATIE has four free parameters (one for each mode: τ, ε, φ, K), which were fitted on the aggregate choice rates (long-term results) of the 60 problems in the Estimation Set of the TPT (Erev et al., 2010), with the restriction that it should capture the qualitative impact curve patterns presented in Fig. 6. Best fit (MSD score of 0.0058) was obtained for $\tau = 0.29$, $\varepsilon = 0.30$, $\varphi = 0.71$, K = 2. The fitted model predicts the aggregate choice rates of the 60 Competition set problems better than the winner in the TPT: Its MSD score on these 60 problems is 0.0056. Appendix B presents the prediction for each of the 120 problems.

The fitted model also captures the short-term phenomena considered above. Its predictions for the mean impact curve in treasure, disaster and non-rare problems of the TPT, are provided in Fig. 6. As can be seen, CATIE nicely captures all major sequential dependencies phenomena observed in the full data and provides excellent quantitative fit.

3.2.1. Alternative models

The predictions of CATIE were compared to the predictions of 13 models proposed for the current data: the three TPT baseline models, eight submissions to the TPT, and to the two best available post hoc models suggested (Lejarraga, Dutt, & Gonzalez, 2012; Shteingart, Neiman, & Loewenstein, 2013). This analysis reveals that none of these 13 models provides better predictions than CATIE for the aggregate choice rates in the Competition set (lower MSD score). The three closest contenders are Lejarraga et al.'s (2012) IBL model (Competition set MSD of 0.0060), Shteingart et al.'s (2013) homogenous resetting of initial conditions (RIC) Q-learning (Competition set MSD of 0.0064), and Erev et al.'s (2010) explorative sampler with recency (the TPT best baseline; Competition MSD of 0.0066). Comparison of CATIE with these three models reveals CATIE not only outperforms them in a prediction of the long-term results, but also has a large advantage over them in capturing the short-term effects of learning. In particular, both IBL and RIC Q-learning predict a strong positive recency effect that is violated by the wavy recency effect and the surprise-triggers-change pattern documented here.

The four best models proposed for the current data, CATIE, IBL, RIC Q-learning, and explorative sampler with recency, all imply reliance on small samples. Yet, the four models differ with respect to the abstraction of the process that implies reliance on small samples. CATIE assumes similarity-based reasoning (where similarity is sequence-based). Explorative sampler with recency assumes reliance on a randomly selected sample (and a bias toward the most recent outcome). Both IBL and RIC Q-learning assume some form of forgetting and thus reliance on the most recent outcomes. The fine quantitative fit of all four models in capturing the long term effects suggest that it may be sufficient to assume reliance on small samples in order to have a good approximation of the aggregate behavior.

3.2.2. The role of inertia and a limitation of CATIE

One interesting observation that emerges from the fitting procedure of CATIE is that it implies very high choice inertia. Specifically, the model implies that if the agent does not make a decision based on either the trend or the exploration modes, the probability that it would simply repeat its last choice is 0.71. Moreover, analysis of the trial-by-trial predictions of CATIE reveals that the model predicts repetition of the last choice in 86% of the trials, suggesting few changes in behavior across the 100 trials of each problem. The data of the TPT corroborates this prediction. Across problems and participants, the mean repetition rate is also 86%. Hence, the high inertia rate CATIE predicts seems like a desired property for a model for the TPT data.

However, this high inertia rate has two important implications. First, it implies "stickiness" of choice, even when the agent happens to choose a clearly inferior option. For example, consider a case in which one option is far better than the other. A high inertia rate implies that if the agent happens to select the inferior option (e.g. because of exploration), the choice of this poor alternative is likely to prolong, possibly for quite a while. Notice this may happen even if the agent has already learnt to prefer the better option, because the model assumes inertia precedes exploitation in the choice process.

A second implication of the high inertia rate involves the case of sequential patterns. Specifically, consider a problem in which the Risky option perfectly alternates between good and bad outcomes. Clearly, the high inertia rate would prevent CATIE from exploiting this pattern fully. Again, this is true even if the model has already "discovered" that the best strategy is to alternate between the options.

To check the importance of these implications, we analyzed another published dataset (Biele et al., 2009), in which there are actual sequential dependencies between the outcomes that one of the options generate. Specifically, in some of the problems of this dataset, one option is likely to alternate between good and bad outcomes. Furthermore, in some of the problems, there is a clearly superior option (with respect to the expected values of the options). Thus, the data presents a particularly difficult challenge for CATIE and the high inertia it implies.

[10] For ease of conceptualization, one may think of CAB-k rules as simple decision trees. For example, a CAB-4 rule defines a tree with four levels of splits. The first split is made according to the relative sign (High or Low) of the most recent outcome, the second split is made according to the relative sign of the second-most recent outcome, and so on. Thus, each route ending in a terminal node (or leaf) of the tree defines a sequence of length four, or a contingency. Each past experience belongs to exactly one such contingency. Averaging the payoffs obtained in all past experiences that belong to the same contingency produces the contingent average of the node. Unlike CAT, designed for with-forgone payoff settings, the sequence of outcomes includes the outcomes observed from either alternative (not only the risky one).

[11] If both contingent averages are equal, the model chooses randomly. Moreover, if the history of experiences is such that no sequence of length k has been observed (i.e. there are no contingencies of length k), then the model chooses the option with the higher average payoff in all previous trials.

The data comes from Experiment 2 in Biele et al.'s (2009) paper. It includes 20 different choice problems similar to those run as part of the TPT. In particular, each problem was a choice between a safe and a binary risky option, run for 100 trials with feedback limited to the obtained outcomes. In contrast to the TPT problems, the probabilities of getting high or low outcomes from the risky option were not fixed. Rather, they were contingent on the outcome in the previous trial: The probability of a high outcome was p if the previous outcome was high and q if the previous outcome was low. Appendix C presents the data in more detail.

Biele et al. (2009) developed a model for this data, and evaluated it based on its ability to capture both the mean aggregate risk-rate and the mean aggregate conditional choice probabilities. Specifically, the model aimed to capture, for each problem: the risk-rate given a safe choice in the previous trial, the risk-rate given risky choice and a high outcome in the previous trial, the risk-rate given a risky choice and a low outcome in the previous trial, and the overall risk rate. Following their procedure, we evaluated CATIE (with the same parameters from the TPT) based on the MSD for each of these four measures. The results show that CATIE does not capture well the results of this experiment. Specifically, the mean of the four MSD scores computed is 0.0243. Conversely, Biele et al.'s 4-mode contingent sampler model (developed to capture this data) obtains much better mean MSD score of 0.0095. Using problem as unit of analysis, the difference of these scores is highly significant, $t(19) = -3.25$, $p = 0.004$. A problem-by-problem examination of CATIE's predictions reveals that the model's largest failures are in problems in which the expected values of the options are very different (i.e. there is a clearly better option), and in problems in which alternation is particularly beneficial (see Appendix C).

3.2.2.1. An extension of CATIE. The analysis above shows CATIE fails to capture behavior in the Biele et al. (2009) data. It also suggests a reason for this failure: The assumption inertia is very likely even if the model "knows" inertia leads to poor performance. In this section, we present a minor extension of CATIE that fixes this issue. We then show this extension captures well behavior both in the TPT data and in Biele et al.'s data.

The extension to CATIE assumes that if the agent is about to choose according to the inertia mode, but the use of this mode is particularly counterproductive, there is probability $\delta \geqslant 0$ (a free parameter) that the agent notices it. If this happens, then inertia is not used, and the next possible mode of CATIE, the exploitation mode (the use of CA rules), is used instead. Note that when $\delta = 0$, this model coincides with CATIE.

To keep track of cases in which inertia is particularly counterproductive, the model maintains, for each possible option, its "average exploitation payoff". This equals the average observed payoff that an option generates in all past trials in which exploitation prescribed choice of that option (whether the option was chosen using the exploitation mode or using another mode). The model then assumes inertia is counterproductive in two cases. The first case is when the average exploitation payoff is higher than both options' observed grand means. That is, the use of the exploitation mode is superior to choosing either of the options consecutively. Notice this is likely to happen when the exploitation mode detects a real pattern and can take advantage of it. Thus, this assumption helps the model avoid inertia when a pattern that may require alternation is detected.

The second case in which inertia is considered particularly counterproductive occurs when two conditions are met. First, the previous choice the model made differs from the choice that it would have made had it followed the exploitation mode in that trial. That is, in the previous trial, the model expected the Risky option to generate the higher payoff, but chose the Safe option, or vice versa. The second condition is that the "average exploitation payoff" of the option that inertia currently implies is sufficiently lower than the "average exploitation payoff" of the alternative option. Here, we define sufficiently lower as a difference that is larger than the mean of the two options' observed standard deviations. Notice that if these two conditions are met, the inertia mode implies repeating a very bad decision that was just made in the previous trial.

This extension of CATIE, with the parameters $\tau = 0.3$, $\varepsilon = 0.48$, $\varphi = 0.78$, $K = 2$, and $\delta = 0.55$, captures the current data far better than the original version of CATIE. Its mean MSD over the four Risky rates is 0.0103 (compared to 0.0243 for the original CATIE). A comparison of this performance with that of the 4-mode contingent sampler model, developed to capture the current data, shows that the two models do not statistically differ in terms of their accuracy, $t(19) = -0.31$, $p = 0.76$.

Moreover, with the same set of parameters, we evaluated the performance of this CATIE extension on the TPT data. The MSD of the TPT estimation set was 0.0057, and the MSD of the TPT competition set was 0.0060. These results are very similar to the results of the original version of CATIE. Furthermore, the aggregate impact curves of this extension, shown in Fig. 6, are very similar to those of CATIE as well. Thus, this extension of CATIE not only captures well behavior in the new data from Biele et al. (2009) sequential dependencies dataset, but also maintains very high accuracy for the TPT data (still better than any other model previously suggested for this data).

3.2.3. Individual choice behavior and a second limitation of CATIE

With the exception of k, the length of the sequence that defines the current contingency, CATIE (as well as its extension) assumes homogeneous parameter values across the population. In the current section, we examine how this assumption influences the accuracy of the model, particularly with respect to its predictions for individual behavior. Note that with respect to individual behavior, differences between CATIE and its extension are minor. Thus, in this section, we focus on the original version of CATIE, but we verified all qualitative results apply to its extension as well.

First, it is worth noting that homogeneous parameters do not necessarily imply homogeneous behavior across participants. Indeed, because the model is stochastic in nature, and because the choices of the agent early on influence the information it has in later trials (as there is no forgone feedback), two agents with the same set of parameters may well exhibit very different behaviors for the same problem. Thus, we first investigate the descriptive power of the model with respect to

the distribution of behavior across the population. To do this, we used CATIE to simulate 1000 virtual agents for each problem of the TPT, and computed the model's predicted distributions of risky choice rates in the population (see Appendix D). We then compared the predicted distributions to the actual distributions observed in the experiment.

The results suggest that CATIE captures quite well the observed risk-rate distributions. Using problem as unit of analysis, the correlation between the standard deviations CATIE predicts and the observed standard deviations is 0.72, 95% CI [0.62, 0.79]. Moreover, as is shown in Appendix D. CATIE correctly predicts the shape of the distribution for the vast majority of problems. In particular, it correctly predicts whether the distribution is skewed (to the left or to the right) or relatively flat. One miss of CATIE appears to be an underestimation of the likelihood that participants would be extreme, exhibiting very low or – more so – very high risk-rates. CATIE rarely predicts such extreme risk rates due to its exploration mode, which implies minimal choice rates for each alternative. One possible solution for this miss is to assume variance (heterogeneity) in the value of ε. Another possible solution is to assume some decision makers do not choose randomly when in exploration mode, but develop a bias towards one of the options.

In addition to CATIE's relatively good predictions for the risk-rate distributions, we examined whether it can predict the distribution in short-term behavior. In particular, we computed for each virtual agent two impact curves, one for treasure problems and one for disaster problems. The results reveal that CATIE captures the proportion of participants exhibiting the qualitative phenomena discussed in Section 2.2.2.2. Specifically, CATIE predicts a pattern consistent with surprise-triggers-change (higher impact at $t + 1$ for disaster problems than for treasure problems) for 66% of the virtual agents. In the TPT data, 93 of 150 participants for whom this measure can be computed exhibit this effect (62%, 95% CI [0.54, 0.70]). In treasure problems, CATIE predicts lower impact 2–4 trials after observing the outcome than 12–14 trials after observing the outcome (the second, increasing, part of the wavy impact, as in Fig. 6, top left panel) for 68% of virtual agents. In the TPT data, 92 of 151 participants for whom this can be computed exhibit this effect (61%, 95% CI [0.53, 0.69]). In disaster problems, CATIE predicts higher impact 2–4 trials after observing the outcome than 12–14 trials after observing the outcome (i.e. no wavy effect, but a decreasing impact, as in Fig. 6, top right panel) for 55% of the virtual agents. In the TPT data, 112 of 198 participants for whom this can be computed exhibit the same pattern (57%, 95% CI [0.49, 0.64]). Thus, in each of the three cases, the 95% confidence interval for the proportion includes CATIE's predicted proportion (and note the TPT results corroborate the conclusions drawn in Section 2.2.2.2).

Clearly, even with homogeneous parameters (except k), CATIE captures quite well the distribution of behavior across participants. Yet, it is still possible that the model's assumption of a mixture over decision modes rather than a mixture over participants is too strong. That is, it is possible that certain decision makers are more likely to behave as if they use a specific mode (e.g. following trends) than others. To examine the evidence for possible individual tendencies to use certain modes, we identified four behavioral measures that each corresponds mainly to one of CATIE's decision modes. We then checked the correlations in these measures over the problems participants faced. The first of the four measures is the rate of choice repetition from trial to trial, corresponding to the inertia mode. The second measure is the absolute difference between the observed choice rate and random choice, corresponding to the exploration mode. The third measure is the proportion of trials in which a choice is consistent with trend-following out of the trials in which the agent observes a trend, a measure that corresponds to the trend (heuristic) mode. The fourth measure is the proportion of trials consistent with underweighting of rare events, corresponding to the exploitation mode (as explained above, the longer the length of the sequence used, the smaller the sample used under this mode and thus the more underweighting of rare events occurs).

To compute the correlations, we first split each set of problems that a participant in the TPT faced to two subsets of six problems each (problems the participant faced 1st, 3rd, 5th, and so on were part of the first set, whereas problems he or she faced 2nd, 4th, 6th, and so one were part of the second set). We then computed for each participant the four behavioral measures described above in each subset of six problems. Finally, for each measure, we computed the correlation between the participants' behavior in one subset of problems and their behavior in the other subset (using participant as a unit of analysis). Any correlations found suggest that participants tend to behave similarly across different problems. The results reveal medium to high correlations in all four measures. The repetition rate correlation is 0.83; the difference from random choice correlation is 0.62; the trend-following correlation is 0.38; and the underweighting of rare events correlation is 0.44. These surprisingly high correlations cannot be predicted by CATIE using homogeneous parameter values. Indeed, this version of the model predicts substantially lower correlations (0 for the trend-following measure, and around 0.25 for the other three measures).

Under the common interpretation of these correlations, they emerge from behavioral tendencies brought from the outside world into the lab. If this interpretation is true, then to characterize each individual's behavior, it may be necessary to fit the model to each individual separately. Unfortunately, for the current data and for the current purposes, such fitting procedure is impractical, as there are simply not enough observations per individual for the fitting procedure to be sound. In particular, one of the most important and novel observations in the current study is the wavy impact curve in treasure problems. This curve relies on the choice rate contingent on a treasure observed n trials back; therefore, to reasonably fit the model to an individual, one must have a sufficient number of trials in which that individual observed a rare treasure. Although each participant in the TPT made 1200 decisions (100 decision in each of 12 problems), the mean number of treasures each participant observed is only around seven. This is because only 36% of the problems are treasure problems, the mean choice rate for the risky option that includes the possible treasure is only 27%, and the mean probability for a treasure in each problem is only 6.5%. Thus, we would essentially estimate four parameters based on seven observations on average.

Moreover, the observed correlations are not necessarily the outcome of stable traits or behavioral tendencies participants bring with them into the lab. They may also emerge if participants do not necessarily treat each problem they face in isolation.

For example, consider a participant who, perhaps due to chance, obtained many high outcomes when choosing the risky option in the first game. On one hand, this would lead to more risky choices in this game. But, it may also lead the participant to "feel lucky" in the next game, which would result in more risky choices in that game as well. Alternatively, if a certain decision mode (e.g. trend seeking) was particularly effective in one game (again, due to chance), this may lead a participant to use this mode more frequently in the next game. That is, a participant may learn that specific decision modes are more or less useful in the current task. Such spillover effects from one game to the next can also lead to the observed correlations.

It is outside the scope of the current paper to decide which of the two possible explanations for the observed correlations, stable behavioral tendencies or spillover effects, is more likely. Nevertheless, some support for the latter explanation is given by the analysis presented in Appendix E. It shows the correlation between the participants' behavior in the very first game they faced and their behavior in each of the other games that followed it, as a function of the temporal distance between the games. The stable behavioral tendencies explanation predicts the correlation should remain the same regardless of the distance between the games (roughly constant). The spillover effects explanation, in contrast, predicts a decrease in the correlation as the distance between the games increases. The results support the latter: they show that in all four measures, the correlations tend to decrease as a function of the distance.

To conclude, CATIE with homogeneous parameters (expect k) captures nicely the distribution of behavior in the population, but it has two limitations. First, it tends to underestimate the likelihood of extreme choice rates, and second, it fails to predict high correlations between participants' tendencies to use each of its modes. Given that CATIE seems like a very good model for the aggregate behavior, a reasonable approach for correcting these limitations is to assume its parameters are not as homogeneous as originally assumed. There are two main ways to think of such a procedure. The first is to assume the values of the parameters differ between individuals but are consistent within individual across games. The second is to assume the parameters can change also within individuals, perhaps as a function of an individual's observed history in the task. We leave a detailed investigation of these two correction options to future research (that would possibly require substantially more data points than the 240,000 points the TPT dataset contains).

4. Discussion

> *"Without waiting, passively, for repetitions to impress or impose regularities upon us, we actively try to impose regularities upon the world"* (Karl Popper)

The original formulation of the law of effect focuses on repeated behavior in the same situation. It states:

> Of several responses made to the same situation, those which are accompanied or closely followed by satisfaction to the animal will, other things being equal, be more firmly connected with the situation, so that, when it recurs, they will be more likely to recur; those which are accompanied or closely followed by discomfort to the animal will, other things being equal, have their connections with that situation weakened, so that, when it recurs, they will be less likely to occur. (Thorndike, 1911, p. 244)

Most previous attempts to quantify this robust observation rest on the implicit assumption that in static experimental tasks, in which agents repeatedly face the same decision, all trials belong to the same situation. Under the law of effect, this "same situation" assumption implies a positive recency effect. In contrast, if Popper's assertion is correct, and we tend to detect patterns even in static settings (when none exists), then participants are likely to believe that the situation in these tasks changes over time (Barron & Yechiam, 2009; Gaissmaier & Schooler, 2008; Restle, 1961; Yu & Cohen, 2009). Our results support this latter assertion. The observed deviations from positive recency suggest that participants behave as if they believe the situation in one trial is different than the situation in subsequent trials and consequently a favorable outcome in that trial does not necessarily increase the likelihood of repeating the reinforced behavior in those subsequent trials.

Thus, we do not take our results as evidence against the law of effect, but as a clarification that its applicability is a function of the organism's understanding and framing of the current "situation". Our favorite framework for this process involves the assertion that decision processes entail similarity judgments (Medin, Goldstone, & Markman, 1995). In particular, we believe agents tend to select the response that worked best in similar situations in the past, a formulation of Skinner's (1953) notion of "contingencies of reinforcements".

One key question is thus how similarity between situations is defined. In our previous work (Plonsky et al., 2015), we suggested that in the context of repeated decisions with complete feedback, sequences of outcomes and local trends serve as useful cues of similarity. The current work extends this idea to settings with incomplete feedback. The robustness of the wavy recency effect of rare events and the fine descriptive accuracy of our similarity-based model, CATIE, imply that using sequential patterns as cues for situation-similarity reflects a robust tendency of human learning. We conjecture that this tendency emerged because in many settings, it leads to near-optimal behavior (Plonsky et al., 2015) and because it facilitates the development of evolutionary advantages like language comprehension and natural events prediction (e.g., Goodman, 2001; Newport & Aslin, 2004; Radinsky & Horvitz, 2013; Saffran, Aslin, & Newport, 1996). Indeed, sequence learning has been postulated to be a core property of multiple diverse brain regions (Goschke & Bolte, 2012).

Although we believe sequential patterns are an important way by which people judge similarity, we clearly do not claim they are the only way. Many other forms of similarity exist, and context probably plays a central role in deciding which of these decision makers use in specific settings. For instance, when environmental information is restricted to the obtained outcomes of previous decisions (like in the current settings), similarity is likely to be judged primarily based on these previous outcomes;

in richer choice settings, similarity comparisons may be based on other environmental cues. Yet, we do not think this means that in richer choice contexts similarity judgments would become so chaotic that behavior would be deemed unpredictable. First, we tend to agree with previous works on judgements of similarity stating that although similarity is context-dependent, it is systematic within context (Medin, Goldstone, & Gentner, 1993). Thus, given a particular context, most agents would likely rely on roughly the same situational cues to find past experiences most similar to the current task. Hence, it should be possible to predict the common behavior in that context. Further research is necessary in order to identify the most common similarity features people rely on in different contexts. Second, if only few past situations are considered similar to the current task, then behavior would likely reflect reliance on small samples of past experiences. Reliance on small samples, in turn, would lead to the long-term learning phenomena like a bias toward underweighting of rare events (Erev & Roth, 2014; Hertwig & Erev, 2009). Importantly, note that many times, sequences of obtained outcomes are available as environmental cues in richer choice settings that include many other cues as well. Thus, if decision makers tend to follow sequences in addition to other environmental cues, then it is quite likely that the number of past experiences that would be considered similar is small.

4.1. The short- and the long-term effects of experience

The current analysis contributes to our understanding of the robustness of three short-term effects of experience, and clarifies the relationship between these three effects and the long-term effects of experience. The first part of our analysis shows that the wavy recency effect can be clearly observed (in treasure problems) even when the feedback available to the learning agents is limited to the outcomes of the selected actions. In addition, the results clarify the significance of the interaction between the wavy recency effect and two related short-term learning effects: the hot-stove effect and surprise-triggers-change. The second part of our analysis shows that a model that abstracts the joint effect of these three short-term phenomena, allows fine prediction of the main long-term effect of experience, and in particular underweighting of rare events. Furthermore, the model clarifies that the fact that there is no clear evidence for a wavy recency effect in disaster problems does not imply a different choice process. The model assumes the use of CA rules in every problem, but the interaction of this process with other elements of the model, leads to different impact curve predictions in different types of problems.

In addition to demonstrating the feasibility of capturing both the short- and long-term effects of experience, the current model, CATIE, suggests that all four effects are associated with a single similarity-based decision process, according to which agents choose the option that worked best in similar past situations. First, underweighting of rare events emerges, according to the model, because the current task appears similar to only few of the previously observed trials, which leads to reliance on small samples, and these do not usually include rare events. Second, because recently observed rare events decrease further the number of past experiences that appear similar to the current task, soon after observing rare events, these rare events are considered even less likely to reappear. Consequently, the wavy recency effect of rare events emerges. Third, since the final decision is correlated with the payoffs generated by the options in similar past situations, in settings without forgone feedback, coincidental bad outcomes in one option are much less likely to be adjusted for in future trials than coincidental good outcomes. Thus, the hot stove effect emerges. Finally, increased exploration emerges, according to our model, following outcomes that diverge significantly than those expected under the assumption that sequences tend to reappear. Thus, surprise-triggers-change emerges.

One advantage of the mechanism we propose here is that it also captures a well-known sensitivity to sequences of outcomes. That is, the same process can account for the basic learning phenomena that emerge in static experimental tasks, as well as for results that show humans (and other animals) respond remarkably well to sequential patterns in a wide variety of tasks (e.g., Clegg, Digirolamo, & Keele, 1998; Gaissmaier & Schooler, 2008; Jones, Curran, Mozer, & Wilder, 2013; Lee, 1971; Lewicki, Czyzewska, & Hoffman, 1987; Restle & Brown, 1970).

One shortcoming of our original model, CATIE, involves the simplification assumption that the probability of using the different modes is solely determined by the level of surprise. Indeed, we show that in settings in which the agent's experience with the task is likely to suggest that choice inertia is a poor decision strategy, CATIE fails to capture behavior. In the extension we propose for CATIE, we remove this simplification assumption and thus enable the model to avoid inertia when using it is clearly sub-optimal. Yet, this extension maintains the assumption that the use of the trend mode is purely probabilistic.

A second shortcoming of the model lies in its inability to predict behavioral consistencies participants exhibit across different games. It is possible that these consistencies are a result of stable traits of participants, in which case a procedure of individual parameter fitting of the model may resolve this shortcoming. However, it is also possible (and more in line with initial analysis of the data) that participants in fact develop preferences for one decision mode or another whilst facing the task. If this is the case, it is unlikely that the current model uncovered the full mechanism underlying the choice of a decision mode. One conjecture is that the choice of a decision mode itself is similarity-based. That is, when having to decide between the different decision modes, agents may recall similar past situations (possibly from previous similar games they faced in the current lab session) and choose to act according to a mode that proved itself useful in these similar past situations. Future research should aim to improve our understanding of the choice among the different decision modes.

4.2. The design of CATIE and the choice process

One particularly intriguing implementation choice of our proposed model is the differentiation between two mechanisms that initially appear to have a similar functionality of outcome dynamics detection. These are the trend seeking mechanism,

aimed to detect local changes in the outcomes, and the CA rules mechanism, aimed to detect global changes in the "situation", or the current state of nature. Interestingly, this dual architecture is reminiscent of recent findings in neurophysiological research. Specifically, Bekinschtein et al. (2009) exposed participants to strings of auditory sounds either of the type xxxxy (four identical sounds followed by a deviant sound), which occurred frequently, or of the type xxxxx (five identical sounds), which occurred rarely. Analysis of brain responses clearly indicated two different levels of neural activity in response to sequential regularity violations: One early, in response to a local deviant sound (i.e. y following a string of xxxx), and the other later, in response to a global deviation from the frequent pattern (i.e. a rare x, rather than a frequent y, following a string of xxxx). Basirat et al. (2014) show similar results for three-month-old infants. These different levels of neural response are commonly assumed to result in two distinct mechanisms. One indication for this is the observation that the response to local violations does not require attention, whereas the response to global violations only occurs when the subject is attentive. Thus, if these results are related to the two mechanisms implemented in CATIE (trend following in response to local outcome changes, and CA rules in response to global sequential patterns), manipulating subjects' attention levels should elicit a different shape of the wavy impact curve. For inattentive subjects, the initial short-term positive recency segment (a result of the trend seeking mechanism) should become more salient, whereas the longer-term negative recency segment (a result of the CA rules) should become less robust.

Another design choice of CATIE that allows for intriguing predictions is the assumption that the surprise mechanism is grounded on similarity comparisons. This assumption provides more intuitive predictions for settings with actual sequential patterns than other surprise mechanisms that are not similarity-based. Consider, for example, the surprise mechanism proposed by Nevo and Erev (2012) for the same data, and a setting in which one alternative alternates between +10 and −10. According to the Nevo and Erev's surprise mechanism, an outcome is more surprising if it differs from the outcome that preceded it. Hence, surprise is greater after a trial in which the alternation persists than after a trial in which it abruptly stops. In contrast, according to the mechanism implemented in CATIE, surprise is minimal if the alternation continues, and increases only if the outcomes suddenly diverge from alternation. Note that CATIE's mechanism makes another interesting prediction: because soon after rare events, rare events are considered less likely to appear than after a sequence of frequent events, surprise should be greater if a rare event soon follows another rare event than if it follows a sequence of frequent events. We plan to test this hypothesis in a future study, measuring surprise with psychophysiological measures.

Another related question is motivated by the abstraction of the inertia and heuristic modes as representing rule-based behavior and the exploitation (contingent average) mode as representing instance-based behavior. As suggested by Logan (1988), instance-based behavior is expected to become more frequent as experience in the task is accumulated; hence we might expect that later in the experiment, behavior would become more consistent with the exploitation mode than with the other modes.

The extension we provide for CATIE implements another relatively unique design choice. This extension requires that the agent would know the choice prescribed by the exploitation decision mode, regardless of whether a choice is made based on this mode or not. Clearly, this appears to be a waste of cognitive resources, particularly considering that the exploitation mode in CATIE requires significant computations. However, there is some evidence to support the assumption that the computations associated with exploitation of dynamic settings may be partially independent of the actual decision made. For example, Brown and Steyvers (2009) show participants are highly sensitive to the dynamics of a changing environment, making quite accurate inferences regarding the generating mechanism. Yet, when the same participants are asked to make decisions in the exact same environment, they behave as if they are quite insensitive to the environmental dynamics. In our terms, these results suggest that on one hand, decision makers may indeed perform the computations necessary for the exploitation mode (thus being able to infer the environmental dynamics), but on the other hand may choose in a manner inconsistent with these computations.

4.3. Top-down or bottom-up process?

Common interpretations of observed sequential dependencies in choice attribute them to a belief-based top-down process. For example, the inclination of decision makers to predict a streak of consecutive outcomes would cease is normally attributed to beliefs regarding the generating mechanism (the gambler's fallacy). We agree that top-down processes can lead to behavior that portrays sequential dependencies. In particular, if decision makers indeed tend to select the option that worked best in similar past situations (our favorite framework of decisions from experience), then clearly beliefs regarding which similarity function is best employed in particular situations play a key role in the decision process. Indeed, an implicit assumption of our model is that agents believe sequences are likely to reappear and it is thus useful to pay attention to the pattern of outcomes.

Nevertheless, we do not explicitly model any specific belief system regarding the generating mechanism. In that sense, our model is somewhat more consistent with a reinforcement-based bottom-up process that learns the dynamics of the environment given the observed outcomes. The model clearly captures well the observed behavior. However, it is still quite possible this behavior is the result of a very different, belief-based process. Specifically, all data we analyze in the current paper comes from studies in which outcomes were generated by a computer, and decision makers may be likely to assume a computer is programmed to generate predictable sequences of outcomes. Yet, it is long known that sequential dependencies emerge even when it should be perfectly clear that the generating mechanism is random. For example, Beach and Swensson (1967) showed that most participants predicting the outcome of a die they role themselves still exhibit sequential

dependencies. To verify this holds for the wavy recency effect as well, we analyzed data from a study (Teodorescu, Amir, & Erev, 2013) in which participants had, in addition to complete feedback from each of the options, a full verbal description of the gambles. Although outcomes here were also generated by a computer, the clear description of the (static) gamble should have made clear that trials are completely independent. Still, our analysis shows that the wavy recency effect replicates in this setting as well. Thus, we believe sequential dependencies can emerge even when decision makers do not have specific prior beliefs regarding the generating mechanism.

The distinction between the belief that sequential patterns are worth attending to and the belief that a specific pattern is built into the generating mechanism is subtle. However, it is related to previous findings that suggest that the learning of statistical relations in the environment requires selective attention, but is otherwise automatic (Turk-Browne, Jungé, & Scholl, 2005). Specifically, for an agent to learn statistical relations among visual stimuli, the agent should not only be exposed to the stimuli, but also be motivated to attend to them over competing stimuli. Yet, once attention is directed towards certain stimuli, the statistical relations among them seem to be implicitly extracted, even when they are presented extremely quickly and when the agent has no particular intention to elicit the statistical structure.

5. Summary and conclusions

In many natural choice settings, decision makers soon discover the outcome of their chosen alternative, while the counterfactual outcomes (the outcomes of the forgone alternatives) are not revealed. In such settings, each decision influences not only the payoff the agent obtains from making the decision but also the information the agent has in future decisions. This added complexity may result in fundamentally different processes than those observed in settings of complete feedback. With this in mind, and given previous work showing rare events have a wavy recency effect on choice, we tried to understand whether a wavy recency effect is robust to feedback type. Analysis of published data from more than 90 choice problems that include rare events show that although limited feedback increases problem complexity (e.g. it encompasses an exploration–exploitation tradeoff that yields a hot stove effect), the wavy pattern is not eliminated.

Moreover, we found that the interactions of the wavy recency effect with other robust properties of basic learning processes, like surprise-triggers-change and a hot-stove effect, lead to qualitatively different response patterns in problems with favorable rare events (treasures) and problems with poor rare events (disasters). Indeed, the fact that only in treasure-type problems, the wavy pattern is conspicuous might explain why it has gone mostly unnoticed in previous research.

Acknowledgments

We thank Kinneret Teodorescu and four anonymous reviewers for useful comments on an earlier version of this manuscript. We also thank Adrian Camilleri and Ben Newell for supplying us with their raw data and allowing us to analyze it. This research was supported by the Israel Science Foundation (Grant Number 1480/13).

Appendix A. The "clicking paradigm, with and without forgone outcomes feedback

	With Forgone Feedback	Without Forgone Feedback
Typical Instructions	The current experiment includes many trials. Your task, in each trial, is to click on one of the two keys presented on the screen. Each click will be followed by the presentation **of both keys' payoffs**. Your payoff for the trial is the payoff of the selected key.	The current experiment includes many trials. Your task, in each trial, is to click on one of the two keys presented on the screen. Each click will be followed by the presentation **of the chosen key's payoff**. Your payoff for the trial is the payoff of the selected key.
Typical Trial		
Initial display (the task)	Click on one of the keys 	Click on one of the keys 
Response		
Feedback		
	You selected Left and your payoff for this trial is 1. Had you selected Right your payoff would have been 0	You selected Left and your payoff for this trial is 1.

Appendix B. Problems of the Technion prediction tournament (Erev et al., 2010)

Problem	Safe	Risky			Risky choice rate	
		High	Prob. high	Low	Observed	CATIE
Estimation set						
1	−0.3	−0.3	0.96	−2.1	0.33	0.38
2	−1	−0.9	0.95	−4.2	0.50	0.49
3	−12.2	−6.3	0.3	−15.2	0.24	0.26
4	−25.6	−10	0.2	−29.2	0.32	0.31
5	−1.9	−1.7	0.9	−3.9	0.45	0.46
6	−6.4	−6.3	0.99	−15.7	0.68	0.70
7	−11.7	−5.6	0.7	−20.2	0.37	0.49
8	−6	−0.7	0.1	−6.5	0.27	0.29
9	−6.1	−5.7	0.95	−16.3	0.43	0.50
10	−1.8	−1.5	0.92	−6.4	0.44	0.45
11	−12.1	−1.2	0.02	−12.3	0.26	0.20
12	−6.4	−5.4	0.94	−16.8	0.55	0.59
13	−9.4	−2	0.05	−10.4	0.11	0.19
14	−15.5	−8.8	0.6	−19.5	0.66	0.54
15	−25.4	−8.9	0.08	−26.3	0.19	0.29
16	−18.7	−7.1	0.07	−19.6	0.34	0.25
17	−23.8	−9.7	0.1	−24.7	0.37	0.32
18	−8.1	−4	0.2	−9.3	0.34	0.27
19	−8.4	−6.5	0.9	−17.5	0.49	0.60
20	−4.5	−4.3	0.6	−16.1	0.08	0.21
21	−4.6	2	0.1	−5.7	0.11	0.23
22	8.7	9.6	0.91	−6.4	0.41	0.42
23	5.6	7.3	0.8	−3.6	0.39	0.35
24	−7.5	9.2	0.05	−9.5	0.08	0.19
25	−6.4	7.4	0.02	−6.6	0.19	0.21
26	−4.9	6.4	0.05	−5.3	0.20	0.26
27	1.2	1.6	0.93	−8.3	0.50	0.45
28	4.6	5.9	0.8	−0.8	0.58	0.39
29	7	7.9	0.92	−2.3	0.51	0.52
30	1.4	3	0.91	−7.7	0.41	0.60
31	6.4	6.7	0.95	−1.8	0.52	0.49
32	5.6	6.7	0.93	−5	0.49	0.57
33	6.8	7.3	0.96	−8.5	0.65	0.54
34	−4.1	1.3	0.05	−4.3	0.30	0.26
35	2.2	3	0.93	−7.2	0.44	0.53
36	−7.9	5	0.08	−9.1	0.09	0.26
37	1.3	2.1	0.8	−8.4	0.28	0.28
38	−5.1	6.7	0.07	−6.2	0.29	0.23
39	−6.9	7.4	0.3	−8.2	0.58	0.48
40	5.9	6	0.98	−1.3	0.61	0.60
41	15.5	18.8	0.8	7.6	0.52	0.53
42	17.1	17.9	0.92	7.2	0.48	0.50
43	9.2	22.9	0.06	9.6	0.88	0.79
44	9.9	10	0.96	1.7	0.56	0.50
45	2.2	2.8	0.8	1	0.48	0.54
46	8	17.1	0.1	6.9	0.32	0.26
47	10.6	24.3	0.04	9.7	0.25	0.21
48	18.1	18.2	0.98	6.9	0.59	0.61
49	9.9	13.4	0.5	3.8	0.13	0.24
50	2.8	5.8	0.04	2.7	0.35	0.25
51	12.8	13.1	0.94	3.8	0.52	0.46
52	0.5	3.5	0.09	0.1	0.26	0.24
53	11.5	25.7	0.1	8.1	0.11	0.20

(*continued on next page*)

Appendix B (*continued*)

Problem	Safe	Risky			Risky choice rate	
		High	Prob. high	Low	Observed	CATIE
54	7	16.5	0.01	6.9	0.18	0.18
55	11	11.4	0.97	1.9	0.66	0.63
56	25.2	26.5	0.94	8.3	0.53	0.56
57	7.9	11.5	0.6	3.7	0.45	0.40
58	20.7	20.8	0.99	8.9	0.63	0.70
59	6	10.1	0.3	4.2	0.32	0.31
60	7.7	8	0.92	0.8	0.44	0.41
Competition set						
61	−21.4	−8.7	0.06	−22.8	0.25	0.22
62	−8.7	−2.2	0.09	−9.6	0.27	0.23
63	−9.5	−2	0.1	−11.2	0.25	0.20
64	−9	−1.4	0.02	−9.1	0.33	0.20
65	−4.7	−0.9	0.07	−4.8	0.37	0.29
66	−6.8	−4.7	0.91	−18.1	0.63	0.61
67	−24.2	−9.7	0.06	−24.8	0.30	0.27
68	−6.4	−5.7	0.96	−20.6	0.66	0.58
69	−18.1	−5.6	0.1	−19.4	0.31	0.28
70	−3.6	−2.5	0.6	−5.5	0.34	0.32
71	−6.6	−5.8	0.97	−16.4	0.61	0.71
72	−15.6	−7.2	0.05	−16.1	0.25	0.23
73	−2	−1.8	0.93	−6.7	0.44	0.45
74	−18	−6.4	0.2	−22.4	0.21	0.24
75	−3.2	−3.3	0.97	−10.5	0.16	0.15
76	−23.5	−9.5	0.1	−24.5	0.39	0.31
77	−3.4	−2.2	0.92	−11.5	0.47	0.60
78	−1.7	−1.4	0.93	−4.7	0.41	0.56
79	−26.3	−8.6	0.1	−26.5	0.49	0.34
80	−20.3	−6.9	0.06	−20.5	0.25	0.29
81	1.7	1.8	0.6	−4.1	0.08	0.21
82	9.1	9	0.97	−6.7	0.14	0.15
83	−2.6	5.5	0.06	−3.4	0.28	0.22
84	0.6	1	0.93	−7.1	0.46	0.46
85	−0.1	3	0.2	−1.3	0.21	0.24
86	−0.9	8.9	0.1	−1.4	0.23	0.33
87	8.5	9.4	0.95	−6.3	0.67	0.57
88	2.7	3.3	0.91	−3.5	0.58	0.49
89	−3.8	5	0.4	−6.9	0.39	0.44
90	−8.4	2.1	0.06	−9.4	0.33	0.22
91	−5.3	0.9	0.2	−5	0.88	0.73
92	−7.6	9.9	0.05	−8.7	0.21	0.23
93	−3	7.7	0.02	−3.1	0.28	0.20
94	2.3	2.5	0.96	−2	0.52	0.58
95	8.2	9.2	0.91	−0.7	0.56	0.52
96	2.9	2.9	0.98	−9.4	0.34	0.29
97	−5.7	2.9	0.05	−6.5	0.30	0.21
98	7.6	7.8	0.99	−9.3	0.62	0.70
99	6.2	6.5	0.8	−4.8	0.32	0.28
100	4.1	5	0.9	−3.8	0.46	0.48
101	19.6	20.1	0.95	6.5	0.50	0.50
102	5.1	5.2	0.5	1.4	0.08	0.19
103	9	12	0.5	2.4	0.17	0.22
104	19.8	20.7	0.9	9.1	0.44	0.43
105	1.6	8.4	0.07	1.2	0.20	0.27
106	12.4	22.6	0.4	7.2	0.41	0.36

Appendix B (*continued*)

Problem	Safe	Risky			Risky choice rate	
		High	Prob. high	Low	Observed	CATIE
107	22.1	23.4	0.93	7.6	0.72	0.54
108	5.9	17.2	0.09	5	0.24	0.29
109	17.7	18.9	0.9	6.7	0.57	0.47
110	4.9	12.8	0.04	4.7	0.26	0.24
111	5.2	19.1	0.03	4.8	0.22	0.22
112	12.1	12.3	0.91	1.3	0.41	0.38
113	6.7	6.8	0.9	3	0.41	0.37
114	11	22.6	0.3	9.2	0.60	0.46
115	1.5	6.4	0.09	0.5	0.28	0.20
116	7.1	15.3	0.06	5.9	0.17	0.19
117	4.7	5.3	0.9	1.5	0.66	0.59
118	12.6	21.9	0.5	8.1	0.47	0.51
119	21.9	27.5	0.7	9.2	0.42	0.37
120	1.1	4.4	0.2	0.7	0.38	0.39

Note. Each problem is a choice between a Safe option (a certain payoff) and a Risky option that provides High with probability Prob. High and Low otherwise. Data is taken from the Technion Prediction Tournament (Erev et al., 2010). CATIE is a Contingent Average, Trend, Inertia, and Exploration model.

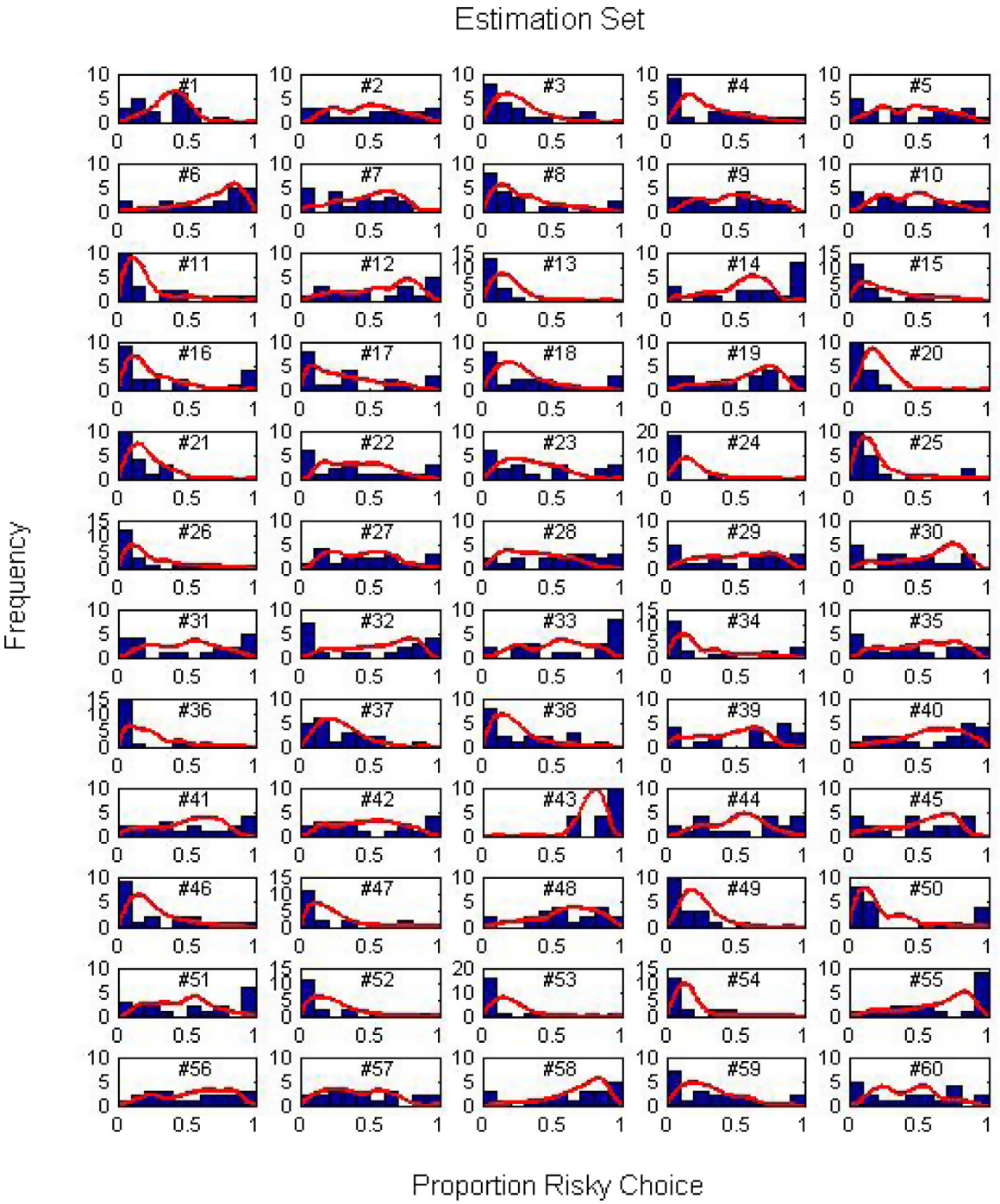
Appendix C. Problems studied in Biele et al.'s (2009) experiment 2

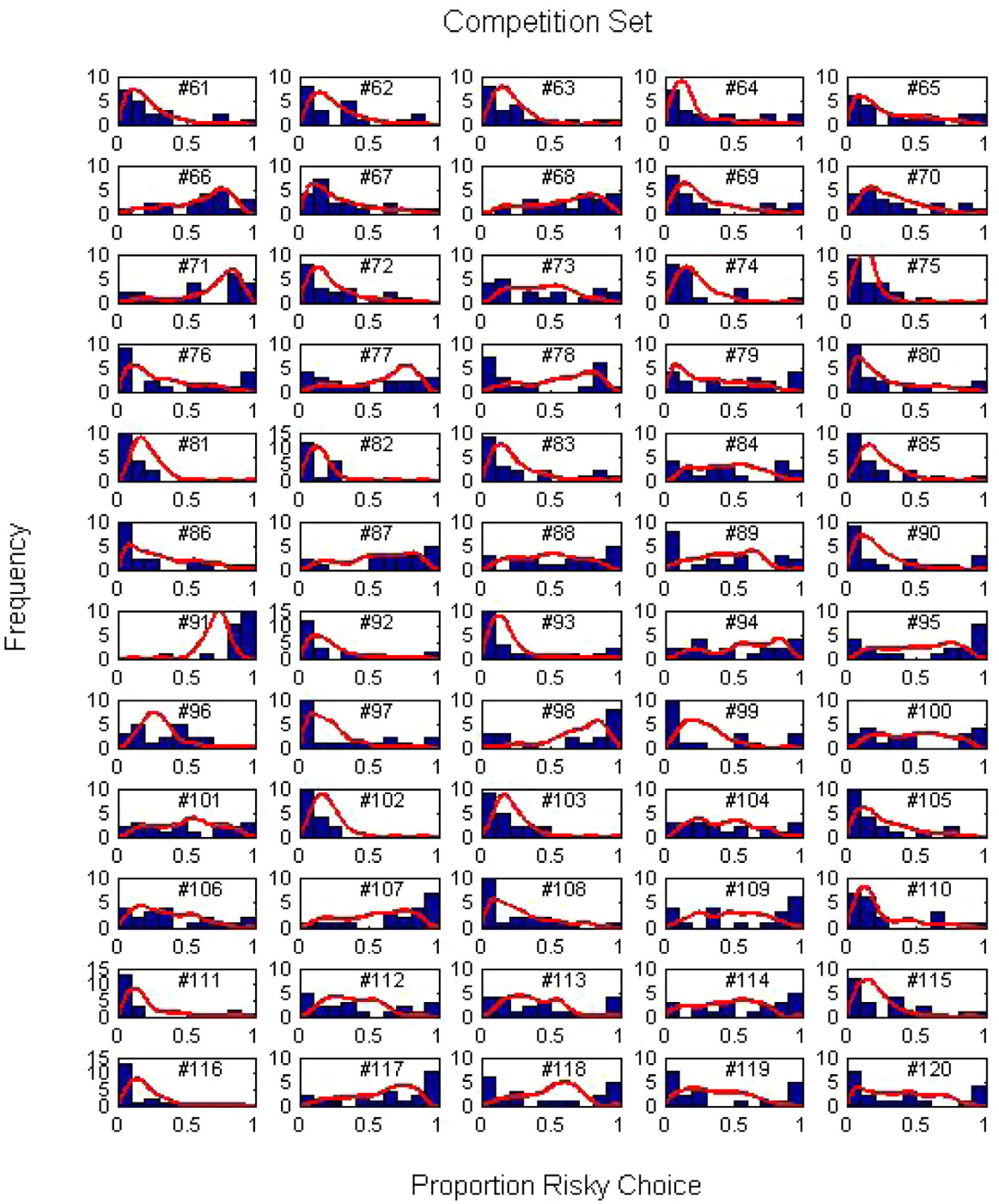
No.	Safe	Risky option				Risky choice rate			Risky after Low			Risky after High			Risky after Safe		
		Low	High	p	q	Obs.	CATIE	Ext.	Obs.	CATIE	Ext.	Obs.	CATIE	Ext.	Obs.	CATIE	Ext.
1	−5	−8	6	0.9	0.5	0.91	0.72	0.77	0.77	0.71	0.64	0.97	0.92	0.88	0.68	0.32	0.55
2	1	−3	6	0.2	0.9	0.46	0.39	0.42	0.58	0.66	0.64	0.55	0.80	0.67	0.37	0.19	0.28
3	−6	−10	0	0.1	0.2	0.22	0.18	0.18	0.42	0.61	0.44	0.80	0.75	0.58	0.15	0.07	0.10
4	−4	−10	−1	0.2	0.9	0.25	0.24	0.27	0.45	0.62	0.58	0.55	0.72	0.62	0.21	0.10	0.14
5	1	−4	3	0.7	0.6	0.37	0.32	0.35	0.54	0.60	0.56	0.68	0.78	0.72	0.21	0.13	0.18
6	7	5	9	0.1	0.2	0.16	0.18	0.17	0.44	0.61	0.42	0.64	0.74	0.56	0.09	0.07	0.10
7	7	3	10	0.8	0.2	0.44	0.34	0.36	0.46	0.64	0.54	0.79	0.81	0.75	0.26	0.13	0.18
8	0	−6	7	0.1	0.3	0.27	0.20	0.20	0.43	0.61	0.48	0.54	0.74	0.59	0.21	0.08	0.11
9	−5	−7	4	0.9	0.7	0.92	0.74	0.78	0.80	0.69	0.64	0.97	0.92	0.88	0.71	0.33	0.56
10	2	−8	5	0.9	0.7	0.71	0.62	0.62	0.59	0.62	0.58	0.83	0.89	0.84	0.47	0.26	0.35
11	−1	−5	10	0.9	0.7	0.95	0.74	0.78	0.88	0.68	0.64	0.98	0.92	0.88	0.71	0.33	0.57
12	7	1	9	0.4	0.1	0.07	0.16	0.15	0.27	0.61	0.40	0.62	0.67	0.48	0.06	0.06	0.09
13	4	−1	6	0.8	0.1	0.21	0.21	0.22	0.40	0.60	0.46	0.79	0.74	0.59	0.10	0.07	0.11
14	−8	−10	9	0.6	0.1	0.36	0.36	0.38	0.53	0.71	0.59	0.82	0.84	0.79	0.26	0.15	0.22
15	−1	−3	10	0.8	0.4	0.84	0.64	0.69	0.79	0.74	0.67	0.98	0.91	0.87	0.58	0.28	0.49
16	−1	−4	2	0.8	0.4	0.73	0.54	0.57	0.67	0.67	0.59	0.90	0.88	0.83	0.58	0.24	0.36
17	0	−3	1	0.1	0.1	0.12	0.16	0.14	0.44	0.60	0.37	0.67	0.74	0.49	0.06	0.06	0.09
18	1	−9	10	0.7	0.7	0.65	0.52	0.53	0.63	0.66	0.61	0.80	0.86	0.81	0.46	0.23	0.31
19	1	0	8	0.2	0.9	0.70	0.52	0.60	0.79	0.69	0.66	0.79	0.87	0.78	0.60	0.26	0.45
20	−6	−9	−5	0.5	0.9	0.25	0.27	0.29	0.44	0.61	0.57	0.56	0.72	0.63	0.19	0.11	0.15

Note. Each problem was a repeated choice for 100 trials between option providing Safe and a Risky option providing either Low or High outcomes. p and q are the probabilities of Risky providing High given it provided High and given it provided Low in the previous trial, respectively. Obs. = Observed choice rate. CATIE is a Contingent Average, Trend, Inertia, and Exploration model. Ext. is an extension of CATIE, described in Section 3.2.2.1. Data is adopted from Biele et al. (2009).

Appendix D. Observed and predicted individual behavior in the TPT

The following figures present, for each problem in the TPT, the observed distribution of behavior across participants (blue histogram), and CATIE's predictions of these distributions (red lines). The number inside each plot is the TPT problem number (see Appendix B). Each predicted distribution was computed by first creating a histogram for the model's simulated predictions, then smoothing the function using spline interpolation, and finally normalizing to scale.


Estimation Set
Frequency
Proportion Risky Choice
#1
#2
#3
#4
#5
#6
#7
#8
#9
#10
#11
#12
#13
#14
#15
#16
#17
#18
#19
#20
#21
#22
#23
#24
#25
#26
#27
#28
#29
#30
#31
#32
#33
#34
#35
#36
#37
#38
#39
#40
#41
#42
#43
#44
#45
#46
#47
#48
#49
#50
#51
#52
#53
#54
#55
#56
#57
#58
#59
#60



Appendix E. Correlations of behavioral measures as a function of temporal distance

Fig. E1.

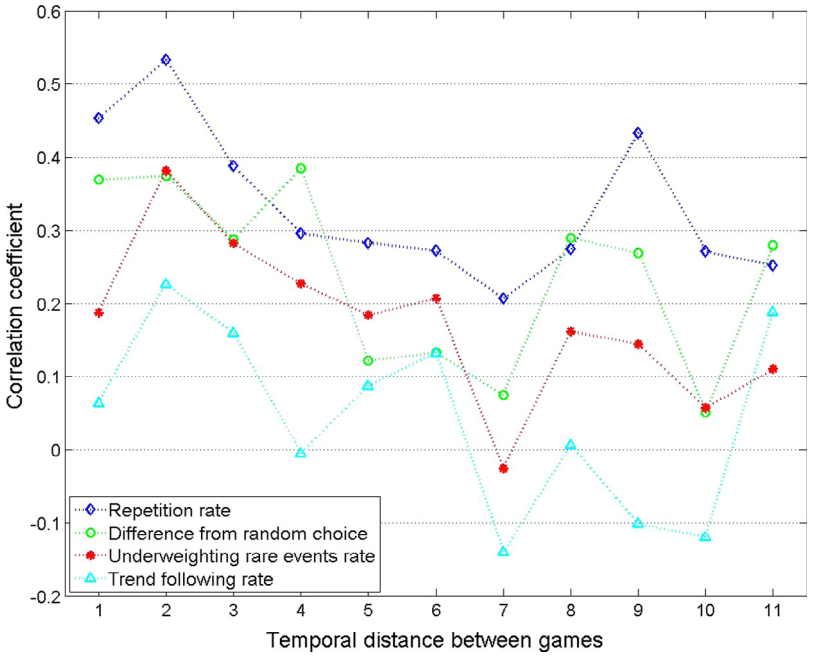


Fig. E1. Correlations of four behavioral measures between first game played and other games played by each participant, as a function of temporal distance between games. Data is taken from the Technion Prediction Tournament (Erev et al., 2010).

Appendix F. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cogpsych.2017.01.002.

References

Anderson, N. H. (1960). Effect of first-order conditional probability in a two-choice learning situation. *Journal of Experimental Psychology, 59*(2), 73–93.

Barron, G., & Yechiam, E. (2009). The coexistence of overestimation and underweighting of rare events and the contingent recency effect. *Judgment and Decision Making, 4*(6), 447–460. <~baron/journal/9729b/jdm9729b.pdf" xlink:type="simple" id="ir010">http://www.sjdm.org/~baron/journal/9729b/jdm9729b.pdf>.

Basirat, A., Dehaene, S., & Dehaene-Lambertz, G. (2014). A hierarchy of cortical responses to sequence violations in three-month-old infants. *Cognition, 132*(2), 137–150. http://dx.doi.org/10.1016/j.cognition.2014.03.013.

Beach, L. R., & Swensson, R. G. (1967). Instructions about randomness and run dependency in two-choice learning. *Journal of Experimental Psychology, 75*(2), 279.

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences, 106*(5), 1672–1677.

Biele, G., Erev, I., & Ert, E. (2009). Learning, risk attitude and hot stoves in restless bandit problems. *Journal of Mathematical Psychology, 53*(3), 155–167. http://dx.doi.org/10.1016/j.jmp.2008.05.006.

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

Busemeyer, J. R., & Myung, I. J. (1992). An adaptive approach to human decision making: Learning theory, decision theory, and human performance. *Journal of Experimental Psychology: General, 121*(2), 177–194. http://dx.doi.org/10.1037/0096-3445.121.2.177.

Bush, R. R., & Mosteller, F. (1955). *Stochastic models for learning*. Oxford, England: John Wiley & Sons, Inc..

Camerer, C. F., & Ho, T.-H. (1999). Experience-weighted attraction learning in normal form games. *Econometrica, 67*(4), 827–874.

Camilleri, A. R., & Newell, B. R. (2011). When and why rare events are underweighted: A direct comparison of the sampling, partial feedback, full feedback and description choice paradigms. *Psychonomic Bulletin & Review, 18*(2), 377–384.

Clegg, B. A., Digirolamo, G. J., & Keele, S. W. (1998). Sequence learning. *Trends in Cognitive Sciences, 2*(8), 275–281. http://dx.doi.org/10.1016/S1364-6613(98)01202-9.

Denrell, J. (2005). Why most people disapprove of me: experience sampling in impression formation. *Psychological Review, 112*(4), 951–978. http://dx.doi.org/10.1037/0033-295X.112.4.951.

Denrell, J. (2007). Adaptive learning and risk taking. *Psychological Review, 114*(1), 177–187. http://dx.doi.org/10.1037/0033-295X.114.1.177.

Denrell, J., & March, J. G. (2001). Adaptation as Information Restriction: The Hot Stove Effect. *Organization Science, 12*(5), 523–538. http://dx.doi.org/10.1287/Orsc.12.5.523.10092.

Di Guida, S., Erev, I., & Marchiori, D. (2015). Cross cultural differences in decisions from experience: Evidence from Denmark, Israel, and Taiwan. *Journal of Economic Psychology, 49*, 47–58. http://dx.doi.org/10.1016/j.joep.2015.04.001.

Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review, 85*(5), 395–416.

Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review, 112*(4), 912–931. http://dx.doi.org/10.1037/0033-295X.112.4.912.

Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S. M., Hau, R., ... Lebiere, C. (2010). A choice prediction competition: Choices from experience and from description. *Journal of Behavioral Decision Making, 23*(1), 15–47. http://dx.doi.org/10.1002/bdm.683.

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review, 88*(4), 848–881.
Erev, I., & Roth, A. E. (1999). On the role of reinforcement learning in experimental games: The cognitive game-theoretic approach.
Erev, I., & Roth, A. E. (2014). Maximization, learning, and economic behavior. *Proceedings of the National Academy of Sciences, 111*(supplement 3), 10818–10825. http://dx.doi.org/10.1073/pnas.1402846111.
Estes, W. K. (1962). Learning Theory. *Annual Review of Psychology, 13*(1), 107–144.
Estes, W. K. (1964). Probability Learning. In A. W. Melton (Ed.), *Categories of human learning* (pp. 89–128). New York, NY: Academic Press.
Fudenberg, D., & Levine, D. K. (1998). *The theory of learning in games. MIT Press books*. Cambridge, MA: MIT Press.
Gaissmaier, W., & Schooler, L. J. (2008). The smart potential behind probability matching. *Cognition, 109*(3), 416–422. http://dx.doi.org/10.1016/j.cognition.2008.09.007.
Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological Review, 118*(4), 523–551. http://dx.doi.org/10.1037/a0024558.
Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science, 27*(4), 591–635. http://dx.doi.org/10.1207/s15516709cog2704_2.
Goodman, J. T. (2001). A bit of progress in language modeling. *Computer Speech & Language, 15*(4), 403–434.
Goschke, T., & Bolte, A. (2012). On the modularity of implicit sequence learning: Independent acquisition of spatial, symbolic, and manual sequences. *Cognitive Psychology, 65*(2), 284–320.
Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science, 15*(8), 534–539. http://dx.doi.org/10.1111/j.0956-7976.2004.00715.x.
Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences, 13*(12), 517–523. http://dx.doi.org/10.1016/j.tics.2009.09.004.
Hohle, S. M., & Teigen, K. H. (2015). Forecasting forecasts: The trend effect. *Judgment and Decision Making, 10*(5), 416–428.
Jarvik, M. E. (1951). Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *Journal of Experimental Psychology, 41*(4), 291–297. http://dx.doi.org/10.1037/h0056878.
Jones, M., Curran, T., Mozer, M. C., & Wilder, M. H. (2013). Sequential effects in response time reveal learning mechanisms and event representations. *Psychological Review, 120*(3), 628–666. http://dx.doi.org/10.1037/a0033180.
Lee, W. (1971). *Decision theory and human behavior*. New York, NY: John Wiley & Sons.
Lejarraga, T., Dutt, V., & Gonzalez, C. (2012). Instance-based Learning: A General Model of Repeated Binary Choice. *Journal of Behavioral Decision Making, 25*(2), 143–153. http://dx.doi.org/10.1002/bdm.722.
Lewicki, P., Czyzewska, M., & Hoffman, H. (1987). Unconscious acquisition of complex procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*(4), 523–530. http://dx.doi.org/10.1037/0278-7393.13.4.523.
Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*(4), 492–527.
Ludvig, E. A., Madan, C. R., & Spetch, M. L. (2014). Extreme outcomes sway risky decisions from experience. *Journal of Behavioral Decision Making, 27*(2), 146–156.
March, J. G. (1996). Learning to be risk averse. *Psychological Review, 103*(2), 309–319. http://dx.doi.org/10.1037/0033-295X.103.2.309.
Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100*(2), 254–278.
Medin, D. L., Goldstone, R. L., & Markman, A. B. (1995). Comparison and choice: Relations between similarity processes and decision processes. *Psychonomic Bulletin & Review, 2*(1), 1–19.
Nevo, I., & Erev, I. (2012). On surprise, change, and the effect of recent outcomes. *Frontiers in Psychology, 3*, 1–9. http://dx.doi.org/10.3389/fpsyg.2012.00024.
Newell, B. R., Rakow, T., Yechiam, E., & Sambur, M. (2015). Rare disaster information can increase risk-taking. *Nature Climate Change*.
Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. *Statistical learning of non-adjacent dependencies. Cognitive Psychology, 48*(2), 127–162.
Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*(1), 1–32. http://dx.doi.org/10.1016/0010-0285(87)90002-8.
Plonsky, O., Teodorescu, K., & Erev, I. (2015). Reliance on small samples, the wavy recency effect, and similarity-based learning. *Psychological Review, 122*(4), 621–647. http://dx.doi.org/10.1037/a0039413.
Radinsky, K., & Horvitz, E. (2013). Mining the web to predict future events. In *Proceedings of the sixth ACM international conference on web search and data mining* (pp. 255–264). New York, NY: ACM.
Rakow, T., & Newell, B. R. (2010). Degrees of uncertainty: An overview and framework for future research on experience-based choice. *Journal of Behavioral Decision Making, 23*(1), 1–14. http://dx.doi.org/10.1002/bdm.681.
Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York, NY: Appleton-Century-Crofts.
Restle, F. (1961). *Psychology of judgment and choice: A theoretical essay*. New York, NY: John Wiley & Sons.
Restle, F., & Brown, E. R. (1970). Serial pattern learning. *Journal of Experimental Psychology, 83*(1, p1), 120–125. http://dx.doi.org/10.1037/h0028530.
Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928. <http://science.sciencemag.org/content/274/5294/1926.abstract>.
Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology, 80*(1), 1–27.
Shteingart, H., Neiman, T., & Loewenstein, Y. (2013). The role of first impression in operant learning. *Journal of Experimental Psychology: General, 142*(2), 476–488. http://dx.doi.org/10.1037/a0029550.
Skinner, B. F. (1953). *Science and human behavior*. New York, NY: Free Press. <http://books.google.com/books?hl=en&lr=&id=Pjjknd1HRElC&oi=fnd&pg=PA1&dq=science+and+human+behavior&ots=iPpkwtE1pI&sig=IP1nEE8_Q-AaQWtavLvDy1uvvII>.
Spiliopoulos, L. (2013). Beyond fictitious play beliefs: Incorporating pattern recognition and similarity matching. *Games and Economic Behavior, 81*(1), 69–85. http://dx.doi.org/10.1016/j.geb.2013.04.005.
Sutton, R. S., & Barto, A. G. (1998). *Introduction to reinforcement learning*. Cambridge, MA: MIT Press.
Teigen, K. H., & Keren, G. (2003). Surprises: Low probabilities or high contrasts? *Cognition, 87*(2), 55–71.
Teodorescu, K., Amir, M., & Erev, I. (2013). The experience-description gap and the role of the inter decision interval. In N. Srinivasan & V. S. C. Pamni (Eds.). *Progress in brain research* (Vol. 202, pp. 99–115). Amsterdam, The Netherlands: Elsevier.
Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Monographs: General and Applied, 2*(4), i-109. http://dx.doi.org/10.1097/00005053-190001000-00013.
Thorndike, E. L. (1911). *Animal intelligence: Experimental studies*. New York: Macmillan.
Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General, 134*(4), 552.
Weiermann, B., & Meier, B. (2012). Incidental sequence learning across the lifespan. *Cognition, 123*(3), 380–391. http://dx.doi.org/10.1016/j.cognition.2012.02.010.
Yechiam, E., & Busemeyer, J. R. (2005). Comparison of basic assumptions embedded in learning models for experience-based decision making. *Psychonomic Bulletin & Review, 12*(3), 387–402. http://dx.doi.org/10.3758/BF03193783.
Yu, A. J., & Cohen, J. D. (2009). Sequential effects: Superstition or rational behavior? In D. Koller, D. Schuurmans, Y. Bengio, & L. Bottou (Eds.). *Advances in neural information processing systems* (Vol. 21, pp. 1873–1880). Berlin: MIT Press. <http://www.pni.princeton.edu/ncc/publications/2009/YuNIPS09.pdf>.