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.