

Recommender Systems and Learning Traps

Ido Erev

Technion-Microsoft Electronic Commerce Research Center, Technion
erev@tx.technion.ac.il

Abstract. Experimental studies of human choice behavior reveal two classes of deviations from optimal choice that should be considered by the designer of recommender systems. The first class of deviations can be described as "presentation effects." In many situations people exhibit high sensitivity to small changes in the presentation of the choice task that do not affect the final outcomes. The second class of deviations can be described as learning traps. In these situations people fail to learn to select the optimal strategies; they appear to converge to inefficient behavior.

Mainstream research in behavioral economics (e.g., Kahneman & Tversky, 1979; Ariely, 2008) tends to focus on presentation effects. The current paper tries to clarify the significance of learning traps. It focuses on two basic properties of decisions from experience that trigger learning traps. The first property, referred to as "underweighting of rare events," is illustrated by the experiments summarized in Figures 1a (results) and 1b (experimental paradigm; see Erev & Haruvy, 2014; Erev & Roth, 2014). In each trial of these experiments the participants are asked to select one of two unmarked keys, and then receive feedback consisting of their obtained payoff (the payoff from the selected key), and the forgone payoff (the payoff that the participant could have received had he selected the other key).

Each participant faced each of the two problems presented in Figure 1a for 100 trials. Both problems focus on choice between a status quo option (0 with certainty) and an action that can lead to positive or negative outcomes. In Problem 1, the action yielded the gamble (-10 with $p = 0.1$; +1 otherwise); this choice has negative expected return ($EV = -0.1$), but it yields the best payoff in 90% of the trials. In Problem 2, the action (+10 with $p = 0.1$; -1 otherwise) has positive expected return ($EV = +0.1$), but it yields the worst payoff in 90% of the trials. The participants received a show up fee of 25 Israeli Shekels (1 Shekel \approx \$0.25) plus the payoff (in Shekels) from one randomly selected trial.

The two curves show the aggregated choice rate of the risky action in 5 blocks of 20 trials over 128 participants that were run in two studies (Nevo & Erev, 2012, Amir et al., 2013). The results reveal that the typical participant favored the risky prospect when it impaired expected return (action rate of 58% in Problem 1 when the EV of the risky prospect is -0.1), but not when it maximizes expected return (action rate of 27% in Problem 2 when the EV of the risky prospect is +0.1). Thus, the typical results

in both problems reflect deviation from maximization. That is, the typical participant behaves "as if" he does not pay enough attention to the rare (10%) outcomes.¹

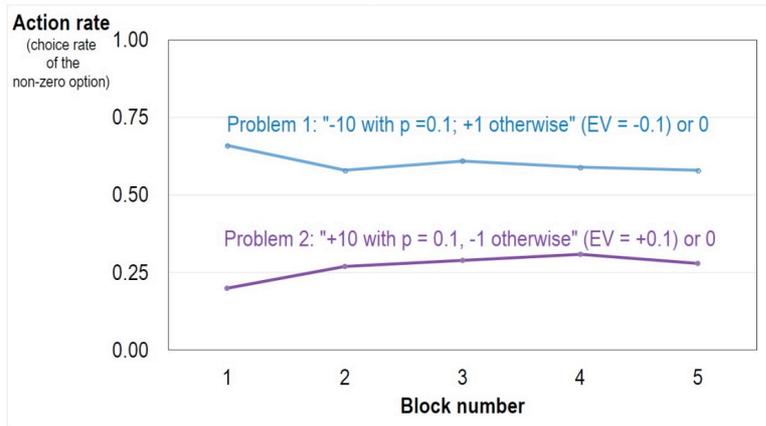


Fig. 1a: Underweighting of rare events. The action rate (proportion of choices of the alternative to the status quo) in the study of Problems 1 and 2 (described in the Figure) in 5 blocks of twenty trials. The curves present the means over the 128 subjects run using the clicking paradigm described in Figure 1b.

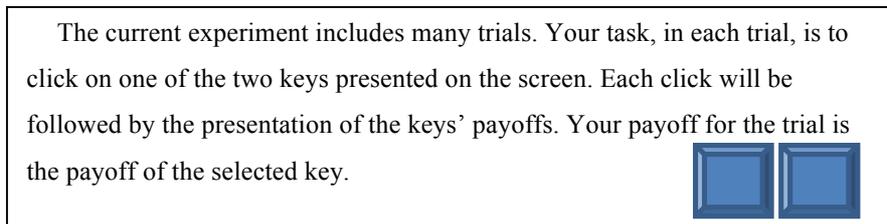


Fig. 1b: The instructions screen in experimental studies that use the basic version of the "clicking paradigm". The participants did not receive a description of the payoff distributions. The feedback after each choice was a draw from each of the two payoff distributions, one for each key.

The studies summarized above focused on situations with complete feedback; the feedback after each trial informed the decision makers of the payoff that they got, and of the payoff that they would have received had they selected a different action. In many natural settings the feedback is limited to the outcome of the selected action, and decision makers have to explore to learn the incentive structure. Analysis of this set of situations highlights the robustness of underweighting of rare events, and shows

¹ Notice that this observation is inconsistent with Prospect Theory (Kahneman & Tversky, 1979). Prospect theory summarizes the results of experiments in which people decide based on a description of the incentive structure, and these studies reveal over-weighting of rare events. The current studies implies that the opposite bias emerge in decisions from experience. That is, the results reflect an experience-description gap (Hertwig & Erev, 2009; Lejarraga & Gonzalez, 2011).

the significance of a second phenomenon: "*the hot stove effect*" (Denrell & March, 2001). When the feedback is limited to the obtained outcome the effect of relatively bad outcomes lasts longer than the effect of good outcomes. The explanation is simple, bad outcomes decrease the probability of repeated choice and, for that reason, they slow reevaluation of the disappointing option. As a result, experience with limited feedback decreases the tendency to select the risky prospect.

Many of the learning traps implied by underweighting of rare events and the hot stove effects can be described as reflections of insufficient exploration. Underweighting of rare event implies that insufficient exploration is particularly likely when the probability of success given exploration is low. In these situations people tend to "give up" too early, and exhibit learned helplessness (Teodorescu & Erev, 2014). For example, they do not learn to use software and applications in the ways that will serve them best, and in some cases they are not be aware of the fact that they are likely to enjoy certain activities (e.g., watching certain type of movies).

The hot stove effect implies that insufficient exploration can also be the product of a random a sequence of bad experiences. For example, a sequence of two bad experiences with a particular product category can lead the agent to stop exploring this category and remember it as unattractive.

Related implications of underweighting of rare events involve a tendency to ignore instructions, sign contracts without reading them, and to skip questionnaires. These behaviors are expected when the extra effort (reading instructions or contracts, and/or filling questionnaires) may be effective in expectation but the common outcome is a waste of time.

Designers of recommender systems can address these and similar learning traps by affecting the incentive structure. Specifically, it is important to understand that in many cases giving the users what they say that they want, may not be good enough. It is possible that the users' behavior reflects a learning trap, and encouraging them to explore and read can help get out from the trap.

References

1. Ariely D (2008) Predictably irrational. New York: HarperCollins.
2. Denrell J, March JG (2001) Adaptation as information restriction: The hot stove effect. *Organization Science*, 12(5), 523-538.
3. Erev I, Haruvy E (in press) Learning and the economics of small decisions. *The Handbook of Experimental Economics*. Princeton University Press. <http://www.utdallas.edu/~eeh017200/papers/LearningChapter.pdf>
4. Erev, I & Roth A. E. (2014). Maximization, Learning and Economic Behavior. *Proceedings and National Academy of Science*, 111, 10818–10825.
5. Hertwig R, Erev I (2009) The description–experience gap in risky choice. *Trends in Cognitive Sciences*, 13, 517-523.
6. Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263-291.
7. Lejarraga T, Gonzalez C (2011) Effects of feedback and complexity on repeated decisions from description. *Organizational Behavior and Human Decision Processes*, 116(2), 286-295.
8. Nevo I, Erev I (2012) On surprise, change, and the effect of recent outcomes. *Frontiers in Cognitive Science*, 3.

9. Teoderescu K, Amir M, Erev I (2013) The experience–description gap and the role of the inter decision interval. In C. Pammi and N. Srinivasan (Eds). *Decision making: neural and behavioural approaches*. Elsevier
10. Teoderescu, Kinneret, and Ido Erev. (2014). Learned helplessness and learned prevalence: Exploring the causal relations among perceived controllability, reward prevalence, and exploration. *Psychological science*. DOI: 10.1177/0956797614543022.