On the description-experience gap, and human reaction to dynamic systems[1]

Ido Erev, Technion

Experimental studies of human decisions in static settings reveal large differences between decisions that are made based on a description of the choice task (as in Kahnman & Tversky, 1979), and decisions that are made based on past experience (Barron & Erev, 2003). The clearest "description-experience gap" (Hertwig & Erev, 2009) involves the weighting of rare events. While the initial reaction to the description of the choice task was found to trigger oversensitivity to rare (low probability) outcomes, the availability of feedback reverses this bias and lead most people to behave as if they believe that "it won't happen to me." The current paper reviews research that documents the description-experience gap, and highlights its implication to the analysis of human reaction to dynamic systems.

The description-experience gap

The top panel in Table 1 summarizes Kahneman and Tversky's study of the impact of rare events in decisions from description. The results reveal oversensitivity to rare events. For example, most participants in their study prefer a "sure loss of 5" over a "1 in 1000 chance to lose 5000." This observation appears to suggest that if our goal is to reduce the frequency of a specific illegal behavior, rare but severe fines (e.g., fine of 5000 for 1 of 1000 of violations) are likely to be more effective than frequent but low fines with the same expected penalty (e.g., a fine of 5 with certainty).

Table 1: Comparison of studies of decisions from description without and with feedback

Study	Main results
Decisions from description (Kahneman & Tversky, 1979) Method: the participants were asked to choose once between the following two hypothetical prospects: S: Sure loss of 5 R: 1 in a 1000 chance to lose 5000; no loss otherwise	Choice rate of option R: 20%. This choice rate suggests that most subjects behave as if the probability of the rare event (-5000) is over-weighted.
The impact of experience (Erev et al., 2017) Method: In each of 25 trials, the participants were asked to choose once between the following prospects. They were paid (in Shekels) for one randomly selected choice, and starting at trial 6, received full feedback (saw the realized payoffs) after each choice. S: Sure loss of 1 R: 1 in a 20 chance to lose 20; no loss otherwise	Initial tendency to choose S (51% before receiving feedback), and a reversal of this tendency after several trials. After 5 trails with feedback, Option R was selected in 65% of the trials.

[1] This text includes paragraphs from Erev et al. (2024) and Erev & Marx (2023)

Subsequent research (Barron & Erev, 2003; Hertwig, Barron, Weber & Erev, 2004) reveals that experience can reverse the impact of rare outcomes. The bottom panel in Table 1 presents one demonstration of this observation. It shows that when people face repeated choices between a "sure loss of 1" and "1 in 20 chance to lose 20" they initially tend to prefer the sure loss, but after less than 5 trials with feedback they change their preference to favor the riskier prospect. Accordingly, the tendency to overweight rare events when considering the initial description is reversed when basing decisions on repeated experiences, leading to under-weighting of rare events in the long run.

Reliance on small samples and the intuitive classifier explanation.

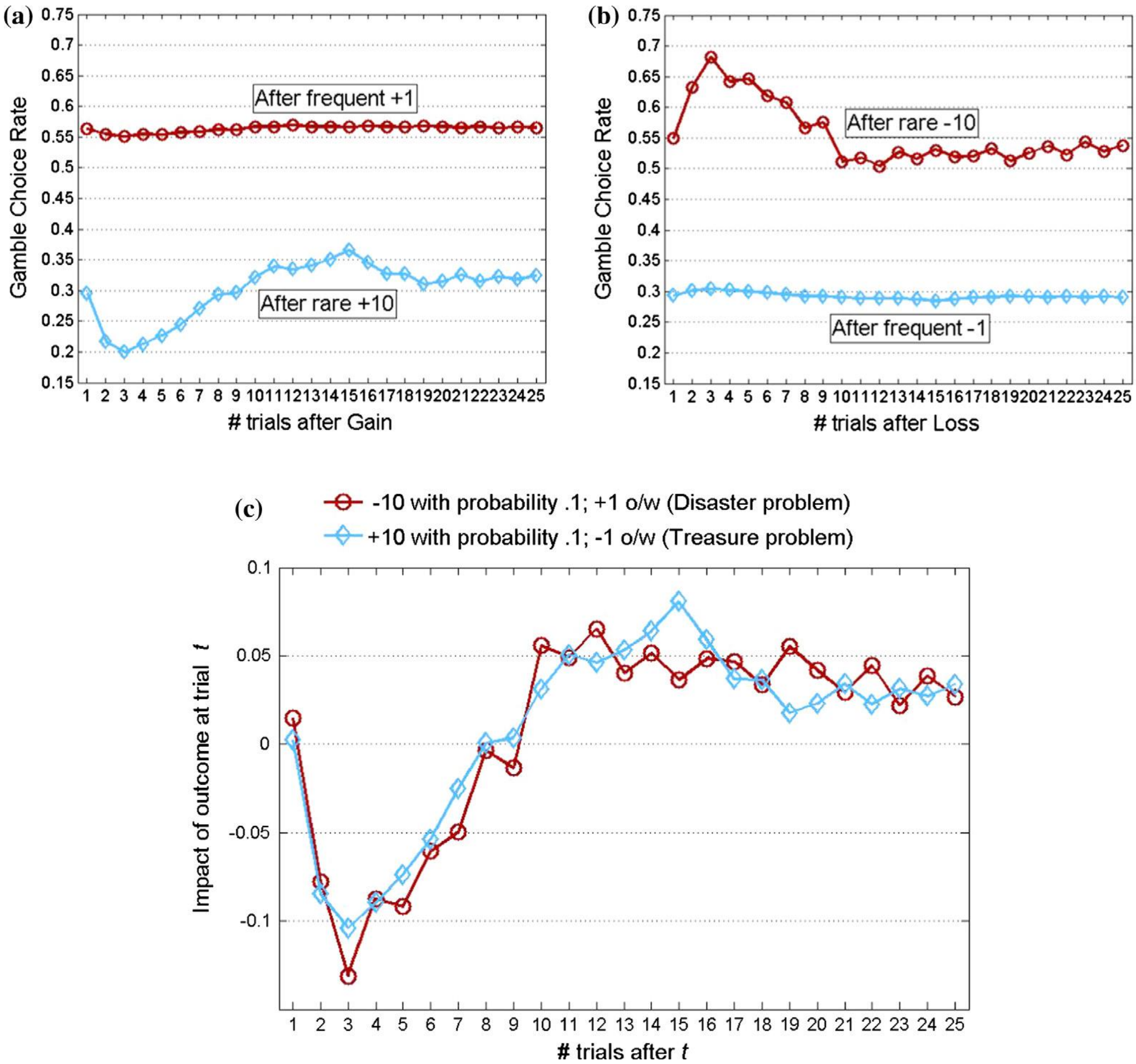
Hertwig et al. (2004) noted that the tendency to underweight rare events in decisions from experience can be captured by assuming that decision makers rely on only small samples of their past experiences. To see why reliance on small samples implies underweighting of rare events, note that the probability that a small sample will not include events that occur with probability $p<0.5$, tend to be larger than 0.5. Specifically, most samples of size k **will not** include a rare event (that occurs with probability p) when the following inequality holds: P(no rare event included) = $(1\text{-}p)^k > .5$. This inequality implies that $k < \log(0.5)/\log(1\text{-}p)$. For example, when $p = 0.05$, $k < 13.51$. That is, when k is 13 or smaller, most samples do not include the rare event (Teodorescu et al., 2013). Therefore, if people draw small samples from the true payoff distributions and choose the option with the higher sample mean, in most cases most of them will choose as if they ignore the possibility that the rare event can actually occur.

The hypothesis that people rely on small samples underlies the most successful models in a series of choice prediction competitions (Erev, Ert, Plonsky, Cohen & Cohen, 2017; Erev, Ert, & Roth, 2010a; Erev, Ert, Roth, et al., 2010b; Plonsky et al., 2019) and can explain many judgement and decision making phenomena (e.g., Erev & Roth, 2014; Erev et al., 2023; Fiedler, 2000; Kareev, 2000; Marchiori et al., 2015).

The wavy recency effect (a violation of the positive recency explanation)

The simplest explanations for the predictive value of models that assume reliance on small samples suggest that it reflects cognitive costs and limitations (see Hertwig & Pleskac, 2010). For example, it is possible that people overweight the easier to remember recent trails, or use a simple "win-stay-lose-shift" heuristic (Nowak & Sigmund, 1993). However, analysis of the sequential dependencies in the data rejects this simple explanation (Plosnky et al., 2015). The clearest evidence against the positive recency explanation comes from studies of decisions made between a safe prospect, and a binary risky prospect with a low probability extreme outcome. The results (see typical findings in Figure 1) reveal a wavy recency effect: The tendency to select the best reply to each occurrence of the rare and extreme outcomes is maximal 11 to 16 trials later. Moreover, the lowest best reply rate was observed 3 trials after the occurrence of the rare, extreme outcome.

Figure 1: Demonstration of the wavy recency effect (adapted from Plonsky and Erev, 2017)



Note: 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. 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 (on the Y-axis) 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).

The intuitive classifiers explanation

Plonsky et al. show that the wavy recency effect, and the descriptive value of the reliance on small samples hypothesis, can be explained with models that share two assumptions: (1) People try to select the option that led to the best outcomes in the most similar past experiences, and (2) The features used to judge similarity include the sequences of recent outcomes. These assumptions imply that the negative recency part of the wavy recency curve (the drop below 0 in Figure 1c) reflects the fact that the number of "similar past experiences" to decisions made immediately after a sequence that includes rare outcomes tends to be small. Table 1 presents examples that clarify this assertion by focusing of the decision in Trial 64 of an experiment that studies the disaster problem of Figure 5. It shows that if the payoff sequence immediately before Trial 64 includes a rare unattractive outcome (loss of -10), agents that select the option that led to the best outcome after a similar sequence are likely to rely on less than 5 past experiences, and are likely to underweight the rare events. Yet, if the sequence of last three recent payoffs does not include a loss, these agents rely on a larger sample (about 44 observations), and are not likely to underweight the rare events.

Table 2: Demonstration of the implications of sequence-based similarity rules

Trials since the last loss	The payoff from the risky option in the three trials before Trial 64			Expected number of similar past experiences in Trial 64	The probability that the average payoff from the risky option over the similar past experiences is positive (and the implied decision reflects underweighting of rare events)
	Trial 61	Trial 62	Trial 63		
More than 3	+1	+1	+1	44.00	0.495
3	-10	+1	+1	4.70	0.593
2	+1	-10	+1	4.79	0.591
1	+1	+1	-10	4.79	0.602

Note: The table considers Trial 64 in the "disaster problem" of Figure 5 ("0 with certainty" or "10% to lose 10, gain of 1 otherwise), assuming that similarity is determined by the three recent payoffs from the risky option. It shows that when the recent payoff sequence includes a rare event, the number of similar past experiences decreases, and the probability of underweighting of the rare event (choosing the risky option) increases.

Plonsky et al. also demonstrate that when the environment is dynamic, judging similarity based on the sequence of recent outcomes can be highly adaptive. For example, consider the thought experiment described in Figure 2. Intuition in this experiment favors a choice of Top in Trial 16. This behavior is implied by the assumption that similarity is determined based on the number of rare and extreme outcomes in the most recent 3 payoffs.

And, under the assumption that the environment is dynamic (e.g., the payoffs are determined by the 4-state Markov chain) it approximates the optimal strategy.

Figure 2: A thought experiment

(a) Task:

In each trial of the current study, you are asked to choose between “Top” and “Bottom”, and earn the payoff that appears on the selected key after your choice is made. The following table summarizes the environment results of the first 15 trials. What would you select in trial 16?

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Top	-1	-1	-1	+2	-1	-1	-1	+2	-1	-1	-1	+2	-1	-1	-1	
Bottom	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

(b) Implications:

In trial 16, intuition favors “Top” despite the fact that the average payoff from “Top” over the 15 trials is negative (-0.4). This intuition suggests that when facing Trial 16, people tend to rely on the most similar previous trials (4, 8 and 12, that like 16 followed a sequence of three -1 outcome). Thus, the choice is made based on only three past experiences.

The static-dynamic puzzle

The analysis presented above suggests that the deviations from optimal decisions in static settings reflects oversensitivity to the possibility that the environment is dynamic. In contrast, experimental studies of decisions in dynamic settings (e.g., Sterman, 1989; Herrnstein, Prelec & Vaughan et al., 1986) document deviations from optimal decisions that suggest insufficient sensitivity to the dynamic nature of the incentive structure. I believe that the study of this puzzle can help facilitate our understanding the human choice behavior.

References

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, *16*(3), 215–233. doi.org/10.1002/bdm.443

Erev, I., Ert, E., Plonsky, O., Cohen, D., & Cohen, O. (2017). From anomalies to forecasts: Toward a descriptive model of decisions under risk, under ambiguity, and from experience. *Psychological Review*, *124*(4), 369–409. https://doi.org/10.1037/rev0000062

Erev, I., Ert, E., Plonsky, O., & Roth, Y. (2023). Contradictory deviations from maximization: Environment-specific biases, or reflections of basic properties of human learning?. *Psychological Review*, *130*(3), 640.

Erev, I., Ert, E., & Roth, A. E. (2010a). A choice prediction competition for market entry games: An introduction. *Games*, *1*(2), 117–136. https://doi.org/10.3390/g1020117

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

Erev, I., Glozman, I., & Hertwig, R. (2008a). What impacts the impact of rare events. *Journal of Risk and Uncertainty*, *36*(2), 153–177. https://doi.org/10.1007/s11166-008-9035-z

Erev, I., Hreib, M., & Teodorescu, K. (2024). Big data without big brothers: The potential of gentle rule enforcement. In *Knowledge and digital technology* (pp. 225-237). Cham: Springer Nature Switzerland.

Erev, I., Plonsky, O. (2023). The J/DM separation paradox and the reliance on small samples hypothesis. In *Sampling in Judgment and Decision Making*, K. Fiedler, P. Juslin and J. Denrell (Eds.)

Fiedler, K. (2000). Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological review*, *107*(4), 659.

Herrnstein, R. J., Prelec, D., & Vaughan Jr, W. (1986). An intra-personal prisoners' dilemma. In *IX Symposium on the Quantitative Analysis of Behavior*.

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.

Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences*, *13*(12), 517–523. https://doi.org/10.1016/j.tics.2009.09.004

Hertwig, R., & Pleskac, T. J. (2010). Decisions from experience: Why small samples? *Cognition*, *115*(2), 225–237. https://doi.org/10.1016/j.cognition.2009.12.009

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292.

Kareev, Y. (2000). Seven (indeed, plus or minus two) and the detection of correlations. *Psychological Review*, *107*(2), 397–402. https://doi.org/10.1037/TO33-295X.107.2.397

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.

Marchiori, D., Di Guida, S., & Erev, I. (2015). Noisy retrieval models of over-and undersensitivity to rare events. *Decision*, *2*(2), 82–106. https://doi.org/10.1037/dec0000023

Nevo, I., & Erev, I. (2012). On surprise, change, and the effect of recent outcomes. *Frontiers in psychology*, *3*, 24. https://doi.org/10.3389/fpsyg.2012.00024

Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-

tat in the Prisoner's Dilemma game. *Nature*, *364*(6432), 56-58.
Plonsky, O., Apel, R., Ert, E., Tennenholtz, M., Bourgin, D., Peterson, J. C., Reichman, D., Griffiths, T. L., Russell, S. J., Carter, E. C., Cavanagh, J. F., & Erev, I. (2019). Predicting human decisions with behavioral theories and machine learning. *ArXiv Preprint ArXiv:1904.06866*.
Plonsky, O., & Erev, I. (2017). Learning in settings with partial feedback and the wavy recency effect of rare events. *Cognitive Psychology, 93*, 18-43. https://doi.org/10.1016/j.cogpsych.2017.01.002
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. https://doi.org/10.1037/a0039413
Sterman, J. (1989) Modeling managerial behavior: Misperceptions of feedback in a dynamicdecision making experiment, *Management Science* 35(3), 321-339.
Teoderescu, K., Amir, M., & Erev, I. (2013). The experience–description gap and the role of the inter decision interval. In V.S. Chandrasekhar Pammi and N. Srinivasan, editors: Progress in Brain Research, Vol. 202, Amsterdam: The Netherlands, 2013, pp. 99-115."