Maximization, learning, and economic behavior

Maximization, learning, and economic behavior

Ido Erev[a,b,1] and Alvin E. Roth[c,1]

[a]Industrial Engineering and Management, Technion, Haifa 32000, Israel; [b]Warwick Business School, The University of Warwick, Coventry CV4 7AL, United Kingdom; and [c]Department of Economics, Stanford University, Stanford, CA 94305-6072

Edited by Brian Skyrms, University of California, Irvine, CA, and approved May 1, 2014 (received for review February 14, 2014)

The rationality assumption that underlies mainstream economic theory has proved to be a useful approximation, despite the fact that systematic violations to its predictions can be found. That is, the assumption of rational behavior is useful in understanding the ways in which many successful economic institutions function, although it is also true that actual human behavior falls systematically short of perfect rationality. We consider a possible explanation of this apparent inconsistency, suggesting that mechanisms that rest on the rationality assumption are likely to be successful when they create an environment in which the behavior they try to facilitate leads to the best payoff for all agents on average, and most of the time. Review of basic learning research suggests that, under these conditions, people quickly learn to maximize expected return. This review also shows that there are many situations in which experience does not increase maximization. In many cases, experience leads people to underweight rare events. In addition, the current paper suggests that it is convenient to distinguish between two behavioral approaches to improve economic analyses. The first, and more conventional approach among behavioral economists and psychologists interested in judgment and decision making, highlights violations of the rational model and proposes descriptive models that capture these violations. The second approach studies human learning to clarify the conditions under which people quickly learn to maximize expected return. The current review highlights one set of conditions of this type and shows how the understanding of these conditions can facilitate market design.

decisions from experience | mechanism design | contingencies of reinforcements | experience–description gap | reinforcement learning

Economics has long used idealized models of perfectly rational individual behavior as useful approximations for explaining market institutions and other economic phenomena. Many of the successful economic inventions (e.g., trading, commodity markets, auctions, matching market institutions) seem to work in ways that are well accounted for by rational models. For example, trading is most likely to be successful when all sides benefit from the trade. Careful empirical studies show, however, important violations of the rationality assumption (e.g., refs. 1–3). Thus, it seems that the usefulness of the leading economic analyses may be enhanced by relaxing the rationality assumption and improving the descriptive value of the underlying models, in part by identifying where the perfect-rationality approximations work more or less well and the purposes for which they may be more or less useful.

Much of what has come to be called behavioral economics can be described as an attempt to make economic analyses more accurate by modifying the rational model to incorporate psychological insights. The most basic rational model, the expected value rule, models people as assigning cash equivalents to possible outcomes, and then selecting the option that maximizes their expected return. The popular generalizations model people as maximizing subjective functions of the objective outcomes; these generalizations fit the observed deviations from rational choice by adding parameters that capture psychological tendencies. Bernoulli's (4) expected utility theory started this influential line of research. It generalizes the expected value rule by adding one psychological parameter: risk aversion or diminishing returns, as axiomatized by von Neumann and Morgenstern (5). Expected utility theory was generalized to subjective expected utility theory by Savage (6) and others. Subsequent modern contributions (e.g., refs. 2 and 7–9) added parameters that capture loss aversion, oversensitivity to rare events, other regarding preferences, and similar psychological tendencies. Gigerenzer and Selten (10) refer to this line of research as the "subjective expected utility repair program." [One shortcoming of this program is the observation that the parameters tend to be situation specific (see refs. 11 and 12). Thus, the derivation of predictions is not always easy.] The main goal of the current paper is to highlight an alternative approach. Instead of testing and repairing the rational model, we focus on studying human learning to clarify the conditions under which people approximately maximize expected return.

The first part of the current paper presents the motivation for our interest in learning, which begins with attempts to understand learning in games, and how the feedback that players get as they learn a game, at the same time as others are learning it, causes behavior in the game to "coevolve," sometimes so that it quickly converges to the predictions reached by models of rational calculation, and sometimes not. What the players know about the game before they begin to play it interacts with how the feedback they get from actually playing it shapes their behavior.

Next, we review learning research into the effect of experience on choice behavior. The results suggest that experience leads to high maximization rates when the optimal choice also provides the best payoff most of the time (under similar settings), but not when this condition is not met. In addition, the results reveal a large difference between decisions that are made based on experience, and decisions that are made based on a description of probabilities and payoffs. People tend to underweight rare events when they rely on experience, and exhibit the opposite bias from maximization when they rely on a description of the choice task.

One interpretation of these observations is that behavior is selected by the contingencies of reinforcements (13), but it is not always possible to detect the correct contingencies. That is, people tend to select the option that led to the best average payoff in similar situations in the past (14, 15), but they are not always good in detecting the truly similar past experiences. Mainstream research in behavioral economics focuses on the conditions that trigger detection errors (overgeneralizations from situations that seem similar but have different optimal strategy), whereas we focus on the conditions that minimize these errors. The clearest members of the class of error-minimizing conditions include repeated tasks in which the optimal choice provides the best payoff most of the time.

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, "In the Light of Evolution VIII: Darwinian Thinking in the Social Sciences," held January 10–11, 2014, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and audio files of most presentations are available on the NAS website at www.nasonline.org/ILE-Darwinian-Thinking.

Author contributions: I.E. and A.E.R. designed research, performed research, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: alroth@stanford.edu or erev@technion.ac.il.

Our summary of the learning literature suggests that novel market designs are likely to be effective when they create a social environment (mechanisms, markets) in which socially desirable behavior provides the best payoff to each individual on average, and most of the time.

Fast and Slow Learning in Games

Experiments reveal that behavior in some games moves quickly toward the predictions of rational choice models, whereas in others it moves much more slowly or not at all in that direction. One example comes from a comparison of the "ultimatum" and the "best-shot" games presented in Table 1. Rational economic theory predicts similar behavior in the two games: the subgame perfect equilibrium predictions give the first mover (player A) a much larger payoff than the second mover (player B), but experimental studies document very different patterns. [Subgame perfect equilibrium (19) is an elegant rational model that implies maximization of expected payoff under the belief that the other agents will maximize their payoff at every opportunity.] In the ultimatum game, the first mover proposes a division of a fixed sum (e.g., $10) between the two players. The perfect equilibrium predicts that the first mover will offer the second mover very little, and that the second mover will accept. However, early play by inexperienced first movers tends to give both players roughly equal payoffs, and this tendency to make high offers persists as the players become more experienced, with second movers showing little change in the propensity to reject low offers that makes moderate behavior wise on the part of the first movers. In the best-shot game, the first mover contributes a quantity x and the second mover contributes a quantity y, which are both inputs into a public good that will equal the maximum (the "best shot") of the two quantities. Both players are charged for the quantity they provide, even though only the maximum contributes to the public good, from which they both earn revenue. With the parameters considered here, the equilibrium predictions are $x = 0$ and $y = 4$. The results show that the first movers quickly learn to select $x = 0$ (that is, to free ride and lay claim to the larger payoff), and the second movers learn to accommodate them. The movement toward the equilibrium prediction in the best-shot game occurs even when the subjects do not receive a description of the game.

In Roth and Erev (ref. 20; and see refs. 21–23), we show how the movement toward equilibrium in one game but not the other fits the predictions of a simple reinforcement learning model, in which players pay attention only to their own payoffs, and tend to repeat actions that have led to good payoffs in the past. In the ultimatum game, there is not much pressure on second movers to learn to accept small offers, because the difference between accepting or rejecting a small offer is small. However, first movers whose offers are rejected (and who therefore earn zero) quickly learn to make offers that will be accepted, because then their payoff is much larger. So in the ultimatum game, first movers learn not to make small offers much faster than second movers learn not to reject them. In contrast, in the best-shot game, first movers do not get much reinforcement from providing a positive quantity of the public good, because the second movers almost uniformly respond, in their best interest, by providing zero themselves. So first movers quickly learn to provide zero, and second movers then learn that, if there is to be any payoff, they must provide it.

Learning in Basic Choice Tasks

To clarify the effect of experience on choice behavior, we focus on experiments that used the simple clicking paradigm described in Fig. 1 (24). In each trial of the experiments described here, 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).

Fig. 2 summarizes the results of two studies that used the clicking paradigm to examine situations in which the prospect that maximizes expected return leads to the worst outcome in most trials. Both studies 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 value (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 ~ $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 five blocks of 20 trials over 128 participants that were run in two studies (25, 26). 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. The participants appear to be risk seekers in problem 1, and risk averse in problem 2. Another way of interpreting the data is that in both cases they reflect underweighting of rare events (22). That is, the typical participant behaves "as if" he does not pay enough attention to the rare (10%) outcomes.

The tendency to underweight rare events in decisions from experience was documented in many settings including: signal detection (27), decisions without forgone payoff (28), one-shot decisions from sampling (29–31), investment decisions (32), market entry games (33), and animal choice behavior (34).

Another set of robust deviations from maximization is illustrated by the four experiments summarized in Fig. 3. The experiments were run using the clicking paradigm and the procedure described above. The results reveal quick learning toward

Table 1. The ultimatum game (16), the best-shot game (17), and the main results observed in an experimental comparison of the two games (18)

	Ultimatum	Best shot
Description of the sequence of play	Stage 1: Player A selects a number ($0 < x < 10$) that represents an offer to player B	Stage 1: Player A selects an integer ($x \geq 0$) that represents A's contribution level
	Stage 2: Player B sees the offer, and then selects between accepting and rejecting it	Stage 2: Player B sees x, and then selects an integer (y) that represents B's contribution
Description of the payoff rule	If B accepts: A earns $10 - x$, B earns x	A earns $R - 0.82x$
	If B rejects: Both players earn 0	B earns $R - 0.82y$
		where $R = Q - \sum_{k=1}^{Q} 0.05(k-1)$, and $Q = \mathrm{Max}(x, y)$
Subgame perfect equilibrium	A: $x = 0$ (or 0.05, minimal offer)	A: $x = 0$
	B: Accept	B: $y = 4$
Main results, mean value of x	Trial 1: 4.2	Trial 1: 1.6
	Trial 10: 4.5 (no change)	Trial 10: 0 (perfect equilibrium)

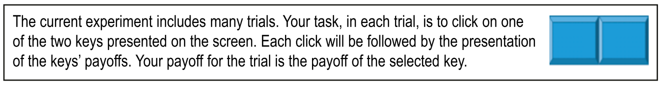

Fig. 1. The instructions screen in experimental studies that use the basic version of the "clicking paradigm." The participants did not receive a description of the payoff distributions. The feedback after each choice was a draw from each of the two payoff distributions, one for each key.

maximization in problems 3 and 4, and almost no learning in problems 5 and 6. Comparison of problems 3 and 5 shows that the added variability (risk) from action in problem 5 reduced the action rate, and comparison of problems 4 and 6 shows the opposite: the added risk increased the action rate. In addition, the results reveals only moderate correlation (0.37) between the risk-taking rates in problems 5 and 6. Busemeyer and Townsend (36) analyzed a similar pattern in a different experimental paradigm and notice that the results reflect a payoff variability effect: High payoff variability moves choice behavior toward random choice, which is related to the observation that variance in outcomes slows learning (37).

Bereby-Meyer and Roth (38) demonstrate the significance of the payoff variability effect in an experiment in which participants repeatedly played 10-period prisoner's dilemma games, i.e., in which they played 10 periods with a fixed other player, and then were matched with a new player to play 10 more periods, until they had played 200 periods, in 20 10-period games. In all conditions of the experiment, they saw what the other player had done after each period. In one condition, the payoffs were deterministic, determined only by what choices the two players had made (e.g., if both players cooperated, they were each credited with \$0.105 for that period). In another condition, the payoffs were stochastic with the same expected value (e.g., if both players cooperated, they each received a lottery that gave them probability 0.105 of receiving \$1.00, and otherwise received nothing). In the deterministic condition, players learned to cooperate in the early periods of their 10-period games, but when payoffs were stochastic, although each player could always see what the other player had done, the variability in payoffs slowed learning to the point that there was little or no learning to cooperate.

Reliance on Small Samples

The deviations from maximization summarized above can be explained if the agents select the option that has led to the best average payoff in the past, and the average payoff is computed based on a small sample (39). For example, if the subjects rely on five past experiences while facing problems 1 and 2, the probability that their sample includes the rare (10%) event is only $1 - (0.9)^5 = 0.41$.

Table 2 summarizes the six experiments described in Figs. 2 and 3 and the prediction of two one-parameter "reliance on small samples" models. The first (sample of 5) assumes random choice at the first trial, and then reliance on a sample of size 5 with replacement from all past experiences. The second model (sample of 9 or less) assumes that the exact sample size is drawn from the set {1, 2, 3…,9}. The results show that both models capture the main results, and that the fit of the second model is better. [Both models are generalizations of the probability matching hypothesis (40). Reliance on sample of size 1 implies probability matching.]

Two-choice prediction competitions (33, 41) demonstrate the value of the assumption that decision makers rely on small samples of past experiences. The best models in both competitions [Stewart and co-workers (41) and Chen et al. (42)] model choices as depending on small samples.

Previous research highlights two possible contributors to the descriptive value of the reliance on small samples assumption. First, it is possible that people rely on small samples to reduce cognitive cost (43–45). A second likely contributor is an attempt to respond to a changing environment, with unobservable changes in the state of nature that can only be inferred from experience (46). Specifically, a decision maker can choose to rely on a small sample because she believes that only a small subset of her past experiences are similar to her current choice task.

The Effect of Limited Feedback

The decisions from experience 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 the phenomena discussed above, underweighting of rare events and the payoff variability effects, and shows the significance of a third phenomenon: the hot-stove effect (47, 48). 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 risky prospects.

The Experience–Description Gap and the Joint Effect of Experience and Description

Comparison of the results summarized in Fig. 2 with mainstream research in behavioral economics reveals an apparent inconsistency. The classical studies that focus on generalizations of the expected value rule suggest that people exhibit overweighting of rare events (2). In contrast, Fig. 2 suggests underweighting of rare events. Recent research suggests that these opposite reactions to rare events are characteristic of an experience–description gap (see review in ref. 49). People tend to exhibit overweighting of rare events when they rely on a description of the probabilities and payoffs, but experience reverses this pattern. Table 3 clarifies this observation. The first two rows (problems 7d and 8d) summarize the classical demonstrations of overweighting of rare events in decisions from description (2). The participants were presented with a description of the payoff distributions (as presented in Table 3) and were asked to make a single decision between the two prospects. The results reveal that the deviation from the expected value rule (which implies indifference between the two prospects) reflect overweighting of the rare (1/1,000) events. People avoid the gamble when the rare event is negative (−5,000 in problem 7d) and prefer it when the rare event is attractive (+5,000 in problem 8d).

The next two rows in Table 3 summarize the results presented in Fig. 2 (problems 1 and 2). These results reveal deviation from the prescription of the expected value rule that involve underweighting of rare (1/10) events. People tend to avoid the gamble when it has positive expected value but the frequent event is unattractive (−1 in problem 1), and prefer the gamble when it

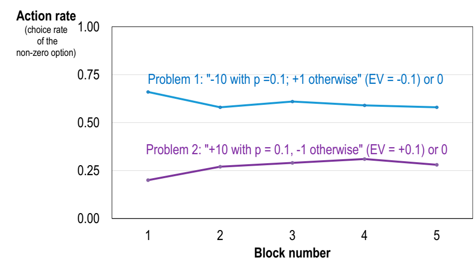

Fig. 2. Underweighting of rare events. The action rate (proportion of choices of the alternative to the status quo) in the study of problems 1 and 2 (described in the figure) in five blocks of 20 trials. The curves present the means over the 128 subjects run (25, 26) using the clicking paradigm.

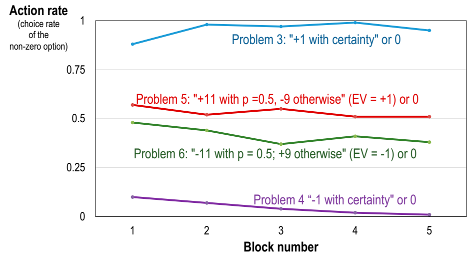

Fig. 3. The payoff variability effect. The action rate (proportion of choices of the alternative to the status quo) in five blocks of 20 trials in the study of problems 3, 4, 5, and 6 (described in the figure). The curves present the means over the 35 subjects run (in refs. 24 and 35) using the clicking paradigm.

has negative expected value but the frequent event is attractive (+1 in problem 2).

The bottom row in Table 3 summarizes the result of a study (ref. 50; and see refs. 51 and 52) that examined the joint effect of description and experience. The participants were presented with the same gamble for 100 trials and could rely on a description of the payoff distributions, and on feedback that was provided after each choice. The results reveal that in the very first trial the typical choice is consistent with overweighting of rare events (55% of the subjects preferred the gamble that promised 64 with probability 1/20 over 3). However, experience reversed their preference. The gamble (that maximizes expected value) was selected only 30% of the time in the last 20 trials. Thus, we see that the effects of description and experience can interact. Another way we can see this is in the experiment of Barron et al. (53), who presented subjects with a choice between two gambles (a sure gain of \$0.10 versus a gain of \$0.13 with probability 0.999 and a loss of \$15 with probability 0.001). One-half of the subjects received this description and then could click on the relevant keys for 100 trials. The other half of the subjects began without any description, as in the clicking paradigm, until they had made 50 choices, at which point the lottery probabilities and payoffs were described to them. So, after the 50th trial, both sets of subjects had exactly the same information (a description plus 50 trials of experience in which they saw the payoffs from both lotteries for 50 periods). However, behavior did not converge between the two sets of subjects: those who had started with a description remained less likely to choose the (higher expected value) risky lottery, with the small chance of a large loss.

A strong effect of description is expected when what players learn is mediated by the behavior of other players who are also learning. This effect is particularly clear in coordination games (see refs. 54 and 55). For example, consider the 5 × 5 Stag Hunt game described in Table 4. The game has two Nash equilibrium points (a choice profile is a Nash equilibrium if no player can benefit from unilateral change of his choice): The first is fair and efficient: both players select E and earn 12. The second is unfair and inefficient: both players select A, one earns 10 and the other 5. A study of 50 trial repeated play of this game with and without description of the payoff rule (56) reveals a clear long-term effect of the initial description. It increased the rate of fair and efficient outcomes in the last 10 trials from 25% to 84%.

Selection by the Contingencies of Reinforcement

Skinner (13) notes that the main results observed in the classical study of animal and human learning can be summarized with the assumption that behavior is selected by the "contingencies of reinforcements." That is, decision makers select the alternative that has led to the best outcomes in similar situations in the past. This summary suggests that the experience–description gap can be a reflection of a difference between two classes of past experiences that decision makers are likely to consider: One class includes old past experiences that occurred before the beginning of the current experiment, and are primed by the description/framing of the task. The second class includes the experiences that occurred during the current task. Initial overweighting of rare events arises when the description leads the decision makers to consider events that are more common (less rare) than the events that determine the outcomes of the current choice. For example, this explanation suggests that the description of problem 8d leads people to rely on past experiences with prizes that occur with higher probability than 1/1,000 (the objective probability to win 5,000 in this problem). Experience increases the proportion of past experiences from the second class (current setting) and for that reason reduces the effect of the description and the implied overweighting of rare events (57).

A more general implication of the Skinnerian summary of the results is the suggestion that it is constructive to distinguish between two classes of factors that are likely to drive choice behavior in a new situation. One class involves the distinct effects of the possible framings and descriptions. These effects are important because they determine which of the decisions makers' old experiences (that occurred before the presentation of the current choice task) will drive initial behavior during the current task. Most research in behavioral economics and social psychology focuses on this wide class of effects. The common findings of these studies suggest deviations from the prediction of the expected value rule. People are sensitive to many factors that do not affect the expected payoff.

A second class involves the possible effects of the new experiences. Understanding of these effects can be used to predict the long-term impact of new incentive structures. The results summarized above suggest that there are many situations in which experience does not lead toward maximization of expected return, and they also suggest that certain conditions facilitate maximization. Specifically, experience leads to a high maximization rate when the optimal choice also provides the best payoff most of the time (under similar settings). The practical implications of this observation are discussed below.

Table 2. The action rate in the first 100 trials of experimental studies of problems 1–6, and the prediction of two one-parameter "reliance on small samples" models

Problem	Action (the alternative to the status quo)	EV	Observed results, %	Model predictions: "Sample of 5," %	Model predictions: "Sample of 9 or less," %
1	(−10, 0.1; +1)	−0.1	58	62	64
2	(+10, 0.1; −1)	+0.1	27	38	36
3	+1 with certainty	+1	95	99	99
4	−1 with certainty	−1	5	1	1
5	(+11, 0.5; −9)	+1	53	50	58
6	(+9, 0.5; −11)	−1	42	50	42

The prospect (x, p; y) pays x with probability p; and y otherwise.

Table 3. Clarification of the experience–description gap

Problem	R rate, %	Proportion of choices consistent with overweighting of rare events, %
Problem 7d: Please choose between: S. −5 with certainty R. (−5,000, 0.001; 0), EV = −5	17	83
Problem 8d: Please choose between: S. +5 with certainty R. (+5,000, 0.001; 0), EV = +5	72	72
Problem 1: Repeated choice from experience between: S. 0 with certainty R. (−10, 0.1; +1), EV = −0.1	58	42
Problem 2: Repeated choice from experience between: S. 0 with certainty R. (+10, 0.1; −1), EV = +0.1	27	27
Problem 9 d and e: Repeated choice based on description and experience between: S. +3 with certainty R. (+64, 0.05; 0), EV = +3.2		
First trial	55	55
After 80 trials with immediate feedback	30	30

The first two rows (problems 7d and 8d, 2) demonstrate overweighting of rare events in decisions from description. The next two rows summarize the indication for underweighting of rare events described in Fig. 2. The last row (50) evaluates the joint effect of description and experience. The results suggest initial overweighting and a reversal of this bias with experience.

Potential Implications

The experimental studies reviewed above highlight conditions under which experience does not guarantee high efficiency, and also suggest that small modifications of the environment can increase efficiency. To clarify the implications of these suggestions to mainstream economic analyses, we consider examples that demonstrate four classes of avoidable inefficiencies outside the laboratory. The first class involves deviations from maximization of expected return. The second class includes inefficiencies that reflect slow learning toward maximization (or toward an efficient Nash equilibrium). A third class involves problems in which insufficient exploration leads to counterproductive self-confirming equilibrium (58). [At a self-confirming equilibrium, the agents select the best option given their experience (and beliefs) and are not aware of the existence of better options.] Finally, we consider situations with multiple Nash equilibria in which experience is likely to lead toward an inefficient equilibrium.

1. Deviations from Maximization. ***Reckless behavior.*** The research reviewed above has two main implications for the design of safe environments (59–61). First is the suggestion that rule enforcement is necessary even when safe behavior (e.g., the use of safety equipment) maximizes the agents' expected return. The explanation of the relevant risks might not be enough. For example, it is not enough to inform people that yearly inspections of their vehicle are beneficial because they reduce the risk of costly mechanical problems and accidents; to facilitate yearly inspections, it is necessary to enforce this behavior. When people make decisions from experience, they are likely to underweight the low-probability–high-hazard events and behave as if they believe "it won't happen to me."

Second, behavior appears to be much more sensitive to the probability than to the magnitude of the punishment. Thus, a gentle rule enforcement policy that employs low punishments with high probability can be very effective (as long as the fine is larger than the benefit from violations of the rule; see ref. 62).

Schurr et al. (61) applied this method in 11 Israeli factories. The basic idea was the design of a mechanism by which supervisors will be encouraged to approach each worker who violates the safety rule and remind him that this behavior might result in injury, and will be recorded (if repeated). The official role of these "violations records" was to allow the management to positively reinforce workers who observe the safety rule by giving these workers a higher probability of winning a lottery. Baseline data were collected about 2 months before intervention. The results reveal a large effect. For example, the proportion of workers who use eye, ear, and hand protections in accordance with the safety rule went up from around 60% to more than 90%. The main increase occurred in the first month of the intervention. More interesting is the observation that the effect of the intervention did not diminish with time.

Warning signals. The gentle rule enforcement policy, described above, enhances safety by changing the problematic incentive structure. A second method to enhance safety focuses on changing the information structure in an attempt to help the agents use the "correct similarity function." For example, safety can be improved by using reliable warning signals. When compliance with the warning signals leads to the best payoff on average and most of the time, the evidence we have discussed suggest that the compliance rate will be high. In fact, it can lead to overcompliance (63). In addition, the decisions from experience literature suggests low compliance rate when the agents learn that the risky choices lead to good outcomes (64).

2. Slow Learning. ***Sealed bid and English auctions.*** The design of auctions has been a very active area of market design (cf. ref. 65). Game-theoretic analysis shows that the assumption that people maximize expected return (and several related assumptions) implies similar revenue to the seller under the most popular

Table 4. A 5 × 5 asymmetric Stag Hunt game

	Player 2				
Player 1	A	B	C	D	E
A	10, 5	9, 0	9, 0	9, 0	9, 0
B	0, 4	0, 0	0, 0	0, 0	0, 0
C	0, 4	0, 0	0, 0	0, 0	0, 0
D	0, 4	0, 0	0, 0	0, 0	0, 0
E	0, 4	0, 0	0, 0	0, 0	12, 12

The availability of a description of the game increases the rate of fair and efficient outcome in the last 10 trials from 25% to 84% (53).

auction mechanisms. In particular, first-price sealed-bid auctions (in which the highest bidder wins, and pays her bid) and second-price sealed-bid auctions (in which the highest bidder wins, but only pays the second highest bid) lead to the same expected outcomes at equilibrium. In addition, when the agents are rational, second-price sealed-bid auctions are equivalent to English (ascending price) auctions.

Learning research suggests that there are also important differences between the popular auctions. Comparison of first- and second-price sealed-bid auctions (66) reveals that second-price auctions create conditions that facilitate learning (the best strategy leads to the best payoff most of time), but first-price sealed-bid auctions do not.

Experimental studies also show faster learning in English (incremental ascending bid) than in second-price sealed-bid auctions (67, 68). Ariely et al. (68) compared these mechanisms in a number of experiments involving the auction of a single item to bidders who knew precisely how valuable it was to them. (In the experiments, what was auctioned was an artificial commodity, and subjects were informed of its cash value to them if they won it, i.e., they were informed that if they won the auction they would receive that cash value minus what they paid to win the auction. The value to each bidder was drawn from a random distribution, which was known to the bidders.) The subjects were given the opportunity to gain experience by participating in auctions over many periods. The payoff maximizing Nash equilibrium prediction in these auctions prescribes bidding the true subjective value of the auctioned good. The results reveal that the modal bid converged to the risk-neutral Nash equilibrium prediction within 15 periods in a second-price auction, and within 9 periods in English auction.

The faster learning in English auction can be the product of the fact that this auction format provides more learning opportunities. In the sealed-bid second-price auction, each period provides only one learning opportunity. For example, a bidder who thought it wise to bid much less than his value, and who lost the auction, only learned this after the auction was over, when it was too late to raise his bid in that auction. In contrast, in the English auctions (in which bidders simultaneously submitted bids, and were informed who was the current high bidder and the current high bid, and had the opportunity to increase their bids if they wished), a bidder who thinks it wise to start with a very low bid learns immediately, and before the auction is over, that a higher bid will be necessary to win.

Notice that, in these experiments, the bidders were given no advice about how to bid. With such advice, they could have learned more quickly to bid their values in the second-price sealed-bid auctions. However, that is because the advice would alert them to something that would be confirmed by their experience. In cases in which the advice is not confirmed by experience, participants often learn to ignore it. In fact, these experiments were motivated by the observation that bidders on the online auction site eBay had learned to "snipe" their bids, i.e., bid in the closing seconds of an auction, despite advice from eBay that they should always bid early (69).

The value of clear instructions in market design. Newly designed markets and allocation mechanisms will begin with all participants inexperienced, at least with respect to the parts of the market that are new. So learning is important in market design, and even markets that move quickly to equilibrium will experience some time away from it. So the behavior of the market while participants are learning, and not just when they are experienced, needs to be a concern.

The changing way that children are assigned to public schools is a case in point. Many American cities give families some choice among the public schools to which their children might be assigned. This assignment was commonly done in a way that made it risky for families to reveal their preferences over schools to the school district. For example, in Boston and many other cities, children would be given priority to attend certain schools (e.g., if they lived close to the school or had an older sibling who attended the school). The school district would try to assign as many families as possible to their first-choice school, and would use the priorities at the school to decide which families would get their first choice if the school was oversubscribed. This meant that a family who was not assigned to the school that they listed as their first choice might find that their second choice was already filled with the children of families who had listed it as their first choice. So a family who had a high priority at a school that was really their second choice might find it too risky to reveal their true preferences, because by doing so they ran the risk of losing the possibility of going to a school that, if they listed it as their first choice, they could be sure of getting. Over the years that such systems were in place, families learned of this risk from experience, and the resulting lessons were passed among families with children entering school, so that eventually many families became quite sophisticated in how they filled out their preference lists for schools, with the consequence that they often settled for schools that they could get rather than revealing to the school district which schools they actually preferred more (70, 71).

Economists have begun to help school districts organize school choice systems that make it safe for families to reveal their true preferences for schools (70, 72). The "deferred acceptance" algorithms by which students are now assigned to schools in a growing number of cities have the property that a student who fails to get into, say, his first-choice school, has just as much chance of getting into his second-choice school as if he had listed it first (73, 74). This is, however, a difficult property of the system to learn from experience. Consequently, it is important to communicate the change in assignment mechanisms by description, and not just by experience, because otherwise the folk wisdom that grew from the experience of previous generations of school children who were assigned under the old mechanism would persist. That is, it is useful not only to design the assignment algorithm so that it makes it a dominant strategy for families to reveal their true preferences, it is important to also explicitly describe this property of the system and give accurate advice about how to use it. Once families are aware that a new mechanism is in place that makes it safe for them to reveal their preferences straightforwardly, it becomes possible to see in the data a change in behavior toward revealing preferences more fully, and this behavior is then reinforced by experience with the assignment mechanism. In other words, we suggest that the basic properties of decisions from experience can help predict when people will trust accurate description of the incentive structure: They are likely to trust when this behavior leads to the best outcome on average and most of the time.

3. Inefficient Self-Confirming Equilibria. ***Corporal punishment.*** Skinner (13) notes that the basic properties of human learning imply that the use of severe punishments in schools can be counterproductive. A negative effect of punishments is likely when the punished agents can respond to the punishment with undesirable avoidance behaviors. For example, punishing pupils for making spelling mistakes is problematic when they can avoid the punishment by stopping coming to school. The hot-stove effect can lead people to select this undesired avoidance strategy even when it impairs their personal long-term well-being. Indeed, the use of punishments in these settings can lead to a counterproductive self-confirming equilibrium (58). Moreover, when the negative outcome of punishments (e.g., avoidance behavior that leads to school dropout) is rare, the tendency to underweight rare events can lead teachers to overuse punishments. Clarification of this helped convinced legislators to ban corporal punishment in schools in many countries.

Insufficient exploration. Corporal punishment is, of course, only one of many factors that can lead to inefficient self-confirming equilibrium. As suggested by the hot-stove effect, there are many situations in which people appear to give up too early. One interesting example is clinical depression. Seligman (75) shows that this disorder can be a result of learned helplessness: a state in which the organism does not explore enough. This interpretation of depression is supported by the observation that cognitive-behavioral therapy, one of the most effective treatments for depression, involves behavioral activation, a procedure in which the therapist encourages patients to participate in activities they no longer engage in, and to try new potentially rewarding activities (76).

Another example is inefficient use of technology and personal abilities. Studies of performance in complex tasks reveal the value of training strategies which enhance exploration of unfamiliar alternatives. One example is the "emphasis change" training protocol (77), according to which the scoring rule is changed on a regular basis, forcing trainees to explore new ways to improve performance. This and similar training strategies were found to enhance performance among pilots (77), basketball and hockey players (see www.intelligym.com), as well as among experimental subjects in a multialternative choice task (78). Teodorescu and Erev (79) show that the conditions under which people exhibit insufficient exploration can be captured with a simple model assuming that the decision to explore is affected by the success of a small set of past exploration efforts in similar situations.

Dinners during conferences. A familiar example of counterproductive but easily avoidable self-confirming equilibria may possibly be observed at dinners during scientific conferences. Many participants come to conferences with the goal of meeting interesting new people, but tend to give up too early. That is, they behave as if early failures to achieve this goal, lead them to believe that this goal is unachievable. One solution to this problem might involve a centralized market for dinners. For example, in a recent conference (http://iew.technion.ac.il/lad/), the dinner parties of the invited speakers were determined by an auction. This method helped the participants meet new people with similar research interests.

4. Coordination Failure in Games with Multiple Equilibria. ***Reducing cheating in examinations.*** Many social problems can be viewed as failures of coordination (55). Consider the problem of cheating during examinations. Problems of this type tend to have two extreme Nash equilibria. In one equilibrium, obeying the (no-cheating) rule is the norm, and the proctors can easily detect and punish deviations if they occur. Thus, no one is motivated to start violating the rule. In a second equilibrium, violation is the norm, and the enforcers are unable to cope with the frequent violations. When the enforcers' resources are limited, the cheating equilibrium can be stochastically stable (80).

The basic properties of decisions from experience imply an easy way to reach the no-cheating equilibrium. It is enough to ensure that each examination will start with a period in which all of the behaviors that appear to reflect cheating will be gently punished. Meeting the eyes of the professor, a comment, or a request to move to the first row could be enough. Once the class reaches the no-cheating equilibrium, maintaining this state should be easy.

Erev et al. (81) evaluate these hypotheses in a field experiment run during final semester examinations of undergraduate courses at the Technion. Traditionally, instructions for examination proctors at the Technion included the following points:

1) The student's ID should be collected at the beginning of the examination;
2) A map of students' seating should be prepared.

Because the collection of the ID is the first step in the construction of the map, the common interpretation of these instructions was that the map should be prepared at the beginning of the examination. Early preparation of the map facilitates deterrence as it signals the possibility of severe punishments (82) that require clear proof of cheating, but it also distracts the proctors, and reduces the probability of gentle punishments at the beginning of the examination.

The experiment compared two conditions. The experimental condition featured a minimal modification of the instructions to proctors that increases the proctors' ability to follow a gentle rule enforcement policy (i.e., promptly warn students whose gaze was wandering). The manipulation was a change of the second instruction to the proctors to the following:

> 2e) "A map of the students seating should be prepared 50 minutes after the beginning of the exam."

Seven undergraduate courses were selected to participate in the study. In all courses, the final examination was conducted in two rooms. One room was randomly assigned to the experimental condition, and the second was assigned to the control condition.

After finishing the examination, students were asked to complete a brief questionnaire in which they were asked to "rate the extent to which students cheated in this exam relative to other exams." The results reveal large and consistent differences between the two conditions. The perceived cheating level was lower in the experimental condition in all seven comparisons.

Traffic jams. Roadway congestion is a source of extreme inefficiency; the estimated cost in the United States for 2005 was \$78 billion (83). Experimental studies reveal quick convergence to inefficient Nash equilibrium in this setting (84). One solution is the use of congestion pricing that builds on our understanding of human adaptation (85). Another solution is the design of attractive alternatives to driving. A promising development involves smart-phone applications that reduce the cost (in terms of waiting time) of cabs and similar services. The current analysis predicts that, when the use of these applications saves time and money, on average and most of the time, they will be highly effective.

Conclusions

The assumption that people always and only maximize expected return is clearly wrong. Nevertheless, many of the successful economic mechanisms rest on this assumption. The current analysis provides a possible explanation for this apparent inconsistency. It suggests that mechanisms that rest on the rationality assumption are likely to be successful when they create an environment under which this assumption is likely to hold. For example, they ensure that the behavior they try to facilitate leads to the best payoff for all agents on average, and most of the time. Basic learning research suggests that, under these conditions, people quickly learn to maximize expected return.

In addition, we suggest that it is convenient to distinguish between two approaches to use behavioral research to improve economic analyses. Mainstream research in behavioral economics tries to highlight violations of the rational model, and propose generalizations of this model that capture these behaviors. The approach discussed here studies human learning to clarify the conditions under which people quickly learn to maximize expected return. It appears to us that this may become an important aspect of our understanding of how the design of economic environments influences the behavior they elicit.

ACKNOWLEDGMENTS. This research benefited from comments from the participants of the Arthur M. Sackler Colloquium of the National Academy of Sciences, "In the Light of Evolution VIII: Darwinian Thinking in the Social Sciences" (January 10–11, 2014, Irvine, CA), and the participants of the Workshop and Winter School on "Learning and Bounded Rationality" (January 2014, Haifa and the Dead Sea, Israel). This research was supported by a grant from the Israel Science Foundation and a grant from the National Science Foundation.

1. Allais M (1953) Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Américaine. *Econometrica* 21(4):503–546.
2. Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47:263–291.
3. Ariely D (2008) *Predictably Irrational* (HarperCollins, New York).
4. Bernoulli D (1738) Specimen theoriae novae de mensura sortis. *Papers Imp Acad Sci St Petersburg* 5:175–192.
5. von Neumann J, Morgenstern O (1947) *Theory of Games and Economic Behavior* (Princeton Univ Press, Princeton), 2nd Ed.
6. Savage LJ (1954) *The Foundations of Statistics* (Wiley, New York).

7. Wakker PP (2010) *Prospect Theory: For Risk and Ambiguity* (Cambridge Univ Press, Cambridge, UK), Vol 44.
8. Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. *Q J Econ* 114(3):817–868.
9. Bolton GE, Ockenfels A (2000) ERC: A theory of equity, reciprocity, and competition. *Am Econ Rev* 90(1):166–193.
10. Gigerenzer G, Selten R (2002) *Bounded Rationality: The Adaptive Toolbox* (Mit Press, Cambridge, MA).
11. Binmore K, Shaked A (2010) Experimental economics: Where next? *J Econ Behav Organ* 73(1):87–100.
12. Ert E, Erev I (2013) On the descriptive value of loss aversion in decisions under risk: Six clarifications. *Judgm Decis Mak* 8:214–235.
13. Skinner BF (1953) *Science and Human Behavior* (Simon and Schuster, New York).
14. Gonzalez C, Lerch JF, Lebiere C (2003) Instance-based learning in dynamic decision making. *Cogn Sci* 27(4):591–635.
15. Gilboa I, Schmeidler D (1995) Case-based decision theory. *Q J Econ* 110(3):605–639.
16. Güth W, Schmittberger R, Schwarze B (1982) An experimental analysis of ultimatum bargaining. *J Econ Behav Organ* 3(4):367–388.
17. Harrison GW, Hirshleifer J (1989) An experimental evaluation of weakest link/best shot models of public goods. *J Polit Econ* 97(1):201–225.
18. Prasnikar V, Roth AE (1992) Considerations of fairness and strategy: Experimental data from sequential games. *Q J Econ* 107(3):865–888.
19. Selten R (1975) Reexamination of the perfectness concept for equilibrium points in extensive games. *Int J Game Theory* 4(1):25–55.
20. Roth AE, Erev I (1995) Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term. *Games Econ Behav* 8:164–212.
21. Gale J, Binmore KG, Samuelson L (1995) Learning to be imperfect: The ultimatum game. *Games Econ Behav* 8(1):56–90.
22. Fudenberg D, Levine DK (1997) Measuring players' losses in experimental games. *Q J Econ* 112(2):507–536.
23. Jehiel P (2005) Analogy-based expectation equilibrium. *J Econ Theory* 123(2):81–104.
24. Erev I, Haruvy E (2013) Learning and the economics of small decisions. *The Handbook of Experimental Economics*, eds Kagel JH, Roth AE (Princeton Univ Press, Princeton), in press. Available at www.utdallas.edu/~eeh017200/papers/LearningChapter.pdf. Accessed May 16, 2014.
25. Nevo I, Erev I (2012) On surprise, change, and the effect of recent outcomes. *Front Cogn Sci* 3:24.
26. Teoderescu K, Amir M, Erev I (2013) The experience–description gap and the role of the inter decision interval. *Decision Making: Neural and Behavioural Approaches*, eds Pammi C, Srinivasan N (Elsevier, Amsterdam).
27. Barron G, Erev I (2003) Small feedback based decisions and their limited correspondence to description based decisions. *J Behav Decis Making* 16:215–233.
28. Barkan R, Zohar D, Erev I (1998) Accidents and decision making under uncertainty: A comparison of four models. *Organ Behav Hum Decis Process* 74(2):118–144.
29. Hertwig R, Barron G, Weber E, Erev I (2004) Decisions from experience and the weighting of rare events. *Psychol Sci* 15(8):534–539.
30. Ungemach C, Chater N, Stewart N (2009) Are probabilities overweighted or underweighted when rare outcomes are experienced (rarely)? *Psychol Sci* 20(4):473–479.
31. Rakow T, Newell BR (2010) Degrees of uncertainty: An overview and framework for future research on experience-based choice. *J Behav Decis Making* 23:1–14.
32. Taleb NN (2007) *The Black Swan: The Impact of the Highly Improbable Fragility* (Random House, New York).
33. Erev I, Ert E, Roth AE (2010) A choice prediction competition for market entry games: An introduction. *Games* 1:117–136.
34. Shafir S, Reich T, Tsur E, Erev I, Lotem A (2008) Perceptual accuracy and conflicting effects of certainty on risk-taking behaviour. *Nature* 453(7197):917–920.
35. Di Guida S, Marchiori D, Erev I (2012) Decisions among defaults and the effect of the option to do nothing. *Econ Lett* 117(3):790–793.
36. Busemeyer JR, Townsend JT (1993) Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychol Rev* 100(3):432–459.
37. Weinstock S (1958) Acquisition and extinction of a partially reinforced running response at a 24-hour intertrial interval. *J Exp Psychol* 56(2):151–158.
38. Bereby-Meyer Y, Roth AE (2006) The speed of learning in noisy games: Partial reinforcement and the sustainability of cooperation. *Am Econ Rev* 96(4):1029–1042.
39. Erev I, Barron G (2005) On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychol Rev* 112(4):912–931.
40. Estes WK (1950) Toward a statistical theory of learning. *Psychol Rev* 57(2):94–107.
41. Erev I, et al. (2010) A choice prediction competition, for choices from experience and from description. *J Behav Decis Making* 23:15–47.
42. Chen W, Liu SY, Chen CH, Lee YS (2011) Bounded memory, inertia, sampling and weighting model for market entry games. *Games* 2(1):187–199.
43. Fiedler K (2000) Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychol Rev* 107(4):659–676.
44. Kareev Y (2000) Seven (indeed, plus or minus two) and the detection of correlations. *Psychol Rev* 107(2):397–402.
45. Hertwig R, Pleskac TJ (2010) Decisions from experience: Why small samples? *Cognition* 115(2):225–237.
46. Plonsky O, Teodorescu K, Erev I (2014) Reliance on small samples, sensitivity to signals and maximization. *Working Paper* (Technion, Haifa, Israel).
47. Denrell J, March JG (2001) Adaptation as information restriction: The hot stove effect. *Organ Sci* 12(5):523–538.
48. Einhorn HJ, Hogarth RM (1978) Confidence in judgment: Persistence of the illusion of validity. *Psychol Rev* 85(5):395.
49. Hertwig R, Erev I (2009) The description-experience gap in risky choice. *Trends Cogn Sci* 13(12):517–523.
50. Lejarraga T, Gonzalez C (2011) Effects of feedback and complexity on repeated decisions from description. *Organ Behav Hum Decis Process* 116(2):286–295.
51. Yechiam E, Barron G, Erev I (2005) The role of personal experience in contributing to different patterns of response to rare terrorist attacks. *J Conflict Resolut* 49(3): 430–439.
52. Jessup RK, Bishara AJ, Busemeyer JR (2008) Feedback produces divergence from prospect theory in descriptive choice. *Psychol Sci* 19(10):1015–1022.
53. Barron G, Leider S, Stack J (2008) The effect of safe experience on a warnings' impact: Sex, drugs, and rock-n-roll. *Organ Behav Hum Decis Process* 106(2):125–142.
54. Battalio R, Samuelson L, Van Huyck J (2001) Optimization incentives and coordination failure in laboratory stag hunt games. *Econometrica* 69(3):749–764.
55. Skyrms B (2004) *The Stag Hunt and the Evolution of Social Structure* (Cambridge Univ Press, Cambridge, UK).
56. Erev I, Grainer B, The 1-800 critique, counter-examples, and the future of behavioral economics. *The Handbook on Methods of Modern Experimental Economics*, eds Frechette G, Schotter A (Oxford Univ Press, Oxford), in press.
57. Marchiori D, Di Guida S, Erev I (2014) Noisy retrieval models of over- and under-sensitivity to rare events. *Working Paper* (University of Southern Denmark, Odense, Denmark).
58. Fudenberg D, Levine DK (1993) Self-confirming equilibrium. *Econometrica* 61(3): 523–545.
59. Erev I, Rodensky D (2004) [Gentle enforcement of safety rules. *A Final Report of a Research Supported by the Committee for Accident Prevention in the Israeli Ministry of Industry & Commerce*] (Technion, Haifa, Israel). Hebrew.
60. Erev I, et al. (2010) The value of gentle enforcement on safe medical procedures. *Qual Saf Health Care* 19(5):1–3.
61. Schurr A, Erev I, Rodensky D (2014) The effect of unpleasant experiences on evaluation and behavior. *Working Paper* (Technion, Haifa, Israel).
62. Gneezy U, Rustichini A (2000) Fine is a price, a. *J Legal Stud* 29:1.
63. Meyer J (2004) Conceptual issues in the study of dynamic hazard warnings. *Hum Factors* 46(2):196–204.
64. Miron-Shatz T, Barron G, Hanoch Y, Gummerum M, Doniger GM (2010) To give or not to give: Parental experience and adherence to the Food and Drug Administration warning about over-the-counter cough and cold medicine usage. *Judgm Decis Mak* 5(6):428–436.
65. Milgrom P (1998) *Putting Auction Theory to Work* (Cambridge Univ Press, Cambridge, UK).
66. Engelbrecht-Wiggans R, Katok E (2007) Regret in auctions: Theory and evidence. *Econ Theory* 33(1):81–101.
67. Kagel JH, Harstad RM, Levin D (1987) Information impact and allocation rules in auctions with affiliated private values: A laboratory study. *Econometrica* 55(6):1275–1304.
68. Ariely D, Ockenfels A, Roth AE (2005) An experimental analysis of ending rules in internet auctions. *RAND J Econ* 36(4):891–908.
69. Ockenfels A, Roth AE (2002) The timing of bids in internet auctions: Market design, bidder behavior, and artificial agents. *AI Mag* 23(3):79.
70. Abdulkadiroğlu A, Pathak PA, Roth AE, Sönmez T (2005) The Boston public school match. *Am Econ Rev* 95(2):368–371.
71. Abdulkadiroğlu A, Sönmez T (2003) School choice: A mechanism design approach. *Am Econ Rev* 93(3):729–747.
72. Abdulkadiroglu A, Pathak PA, Roth AE (2009) *Strategy-Proofness Versus Efficiency in Matching with Indifferences: Redesigning the New York City High School Match* (National Bureau of Economic Research, Cambridge, MA), No. w14864.
73. Gale D, Shapley LS (1962) College admissions and the stability of marriage. *Am Math Mon* 69(1):9–15.
74. Roth AE (2008) What have we learned from market design? *Econ J* 118(527):285–310.
75. Seligman ME (1972) Learned helplessness. *Annu Rev Med* 23:407–412.
76. Beck AT, Rush AJ, Shaw BF, Emery G (1979) *Cognitive Therapy of Depression* (Guilford, New York).
77. Gopher D, Weil M, Bareket T (1994) Transfer of skill from a computer game trainer to flight. *Hum Factors* 36(3):387–405.
78. Yechiam E, Erev I, Gopher D (2001) On the potential value and limitations of emphasis change and other exploration-enhancing training methods. *J Exp Psychol Appl* 7(4): 277–285.
79. Teodorescu K, Erev I (2014) On the decision to explore new alternatives: The coexistence of under- and over-exploration. *J Behav Decis Making* 27(2):109–123.
80. Young HP (1993) The evolution of conventions. *Econometrica* 61(1):57–84.
81. Erev I, Ingram P, Raz O, Shany D (2010) Continuous punishment and the potential of gentle rule enforcement. *Behav Processes* 84(1):366–371.
82. Becker GS (1968) Crime and punishment: An economic approach. *J Polit Econ* 76: 169–217.
83. Schrank DL, Lomax TJ (2007) *The 2007 Urban Mobility Report* (Texas A&M Transportation Institute, College Station, TX).
84. Rapoport A, Seale DA, Erev I, Sundali JA (1998) Equilibrium play in large group market entry games. *Manage Sci* 44(1):119–141.
85. Sandholm WH (2002) Evolutionary implementation and congestion pricing. *Rev Econ Stud* 69(3):667–689.