10

Learning and the Economics of Small Decisions

Ido Erev and Ernan Haruvy

INTRODUCTION

Mainstream analysis of economic behavior assumes that economic incentives can shape behavior even when individual agents have limited understanding of the environment (see related arguments in Nash[1] 1950; Smith[2] 1962). The shaping process in these cases is indirect: The economic incentives determine the agents' experience, and this experience in turn drives future behavior. Consider, for example, an agent that has to decide whether to cross the road at a particular location and time. The agent (say a chicken) is not likely to understand the exact incentive structure and compute the implied equilibria. Rather, the agent is likely to rely on experience with similar situations. The economic environment shapes this decision because it determines the relevant experience.

The current chapter reviews experimental studies that examine this shaping process. In order to clarify the relationship of the research reviewed here to classical research in behavioral and experimental economics it is constructive to consider the distinction between "decisions from description" and "decisions from experience" (Hertwig et al. 2004) exemplified in Figure 10.1. Classical studies in behavioral economics tend to focus on decisions from description: They examine how people decide when they can rely on a complete description of the incentive structure. In contrast, the research reviewed here focuses on decisions from experience. In a pure decision from an experience task (like the one demonstrated in Figure 10.1), the decision makers do not receive a prior description of the incentive structure. Rather, they have to rely on past experience, and gain relevant experience in the course of the experiment.

The two lines of decision research have similar goals but take very different routes towards achieving these goals. As a result, the two routes often identify and focus on different behavioral regularities. The main difference between the two routes is reflected by the relationship of the two lines of research to rational economic theory. The classical studies of decisions from description were designed to test the rationality assumption. The most influential papers in that research stream (e.g., Allais 1953; Kahneman and Tversky 1979; Fehr and Schmidt 1999; Bolton and Ockenfels 2000) present interesting

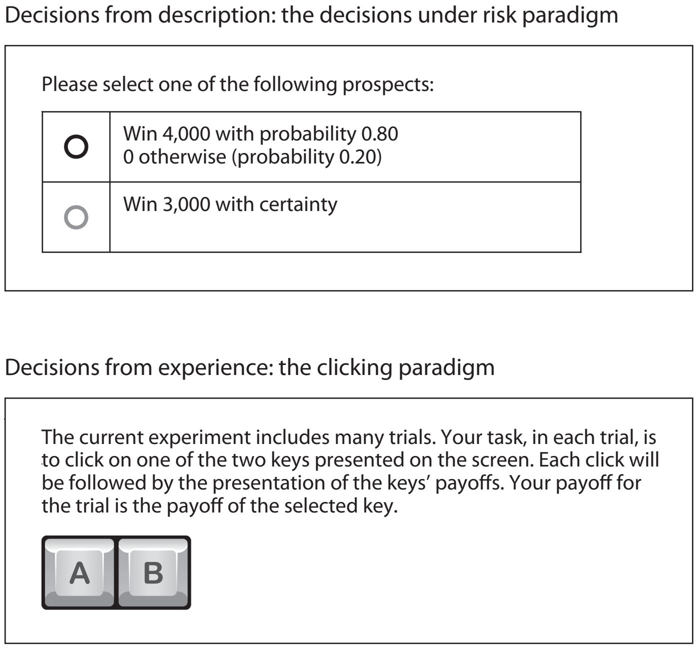

Figure 10.1: Typical instruction screen in studies of decisions from description (using the "decisions under risk paradigm") and studies of decisions from experience (using the "clicking paradigm").

deviations from rational choice and elegant refinements of the rational models that capture these deviations. Gigerenzer and Selten (2001) broadly refer to this line of research as the "subjective expected utility repair program." In contrast, the studies of decisions from experience focus on situations for which rational decision theory does not have clear predictions. When decision makers rely on past experience, almost any behavior could be justified as rational given their experience and the beliefs this fosters. Thus, the study of decisions from experience is not designed to test or refine rational decisions theory; rather, it is intended to expand the set of situations that can be addressed with economic models that provide clear and useful predictions.

The significance of the difference between the behavioral regularities discovered in the two lines of decision research is demonstrated by the effect of rare (low-probability) events. Experimental studies reveal that people exhibit oversensitivity to rare events in decisions from description (Kahnemna and Tversky 1979) and the opposite bias when they rely on experience (see Barron and Erev (2003) and Section 1.1.3). This "experience-description gap" suggests that the common efforts to use models that were calibrated to capture decisions from description in order to address decisions from experience can lead to mismatched conclusions.

Many natural decision problems fall in between decisions from description and decisions from experience. For example, in 2003 when the president of the United States, George W. Bush, had to decide whether or not to engage militarily in Iraq, he could rely on a description of the incentive structure, prepared by his consultants, but he could also rely on historical experiences in similar situations. And it is possible that these experiences could suggest that the description can be biased.

The importance of past experience is particularly clear in the context of small decisions. Small-decision problems are defined here as situations in which the performance of a task requires decisions, and the expected consequences of each decision are relatively small. Many natural activities involve small decisions. For example, the road-crossing task, described earlier, implies several small decisions. The agent can choose whether to start crossing in several points in time and can then choose to change his or her mind.

We believe that small decisions can be of large economic importance. In many cases, small decisions can be highly consequential in the aggregate, and they can also be consequential in some rare specific cases. For example, small driving-related decisions lead to traffic jams, costly accidents, injuries, and even fatalities. Moreover, in many cases small decisions shape subsequent big decisions. For instance, the small decisions between doing homework or watching TV as a child can affect the available alternatives in the big decisions among different career paths. Similarly, the big decision between different investment portfolios is made only if the agent has made the small decision to spend time (at a particular point in time) on evaluating his or her investments.[3]

Economics, psychology, and the clicking paradigm. Most of the studies of learning and decisions from experience were conducted by psychologists and were not designed to evaluate the effect of the quantitative features of the incentive structure; they typically used nonmonetary reinforcements like food, electric shocks, unpleasant noises, and verbal recognition. In order to clarify the economic implications of these studies, we try to replicate the main results using the clicking paradigm presented in Figure 10.1. As demonstrated in Figure 10.1, the clicking paradigm focuses on the effect of experiencing monetary payoffs. The subjects do not receive prior description of the incentive structure and have to base their decisions on the feedback (observed monetary outcomes) of previous decisions. To facilitate evaluation of the effect of this experience, each experiment includes many trials.

In order to illustrate the relationship of the current replications to the original demonstrations of the classical phenomena, we start the discussion of the key phenomena with a description of the original studies. Yet, we pay greater attention to the clicking paradigm.

Another advantage of the clicking paradigm replications involves the standardization of the experimental conditions (Hertwig and Ortmann 2002). For example, it allows the use of the same instructions, same experimental design, and same subject population in the replications of the distinct regularities.[4] Since we focus on phenomena that were already documented in a wide set of conditions with a wide set of subject populations (including very different animals), the focused standardization should not impair external validity. The standardization is expected to clarify the role of the key factor—the effect of experiencing the incentive structure—and facilitate the development of models that capture this effect.

Three cognitive factors and the organization of the current review. Decisions from experience are likely to be affected by three classes of cognitive factors (see Erev and Roth 1999). The first factor involves the cognitive strategies considered by the agents, that is, the strategies from which the agents learn. The cognitive strategies include the possible actions (stage game strategies, "Select the left Key" or "Select the right key" in the basic clicking paradigm), but can also include other strategies like "Try to reciprocate" (see Section 4.3) or "Select best reply to the instructions" (see Section 1.1.9). The second factor involves the exploration policy—that is, the trade-off between collecting information and using the available information in order to get the best outcomes (see March 1991). The third factor is the choice rule: the evaluation of past experiences that determines which strategy is preferred.

We believe that all three factors are important, but we also think that better understanding of the choice rule is likely to be most useful. Our belief is based on the observation that the cognitive strategies and the exploration policy tend to be situation specific. Small changes in the environment can change the strategies available and considered by the agents and can change the value of exploration. In contrast, it is possible that the choice rule reflects more robust properties of the underlying cognitive processes that are likely to be stable over situations and maybe also over species.

This belief led us to start the current review with a focus on phenomena that can be replicated even when the effect of the first two factors is minimized. Specifically, we start with a focus on situations in which (1) it is reasonable to assume that the strategies considered by the agents can be approximated by the possible actions, and (2) exploration does not add information. The most important part of the current review is Section 1.1, which presents six robust behavioral phenomena that emerge in this setting and a simple model that summarizes them. We consider situations in which exploration is important in Section 1.2 and delay the discussion of situations in which more sophisticated strategies are likely to be important to Sections 2, 3, and 4.

Section 2 reviews studies of learning in dynamic environments, and Section 3 reviews studies of learning among of large number of alternatives. The results highlight interactions between the basic properties of learning, summarized in Section 1, and other factors that can be abstracted as "cognitive strategies" that are implied by the task.

Section 4 reviews studies that examine the effect of social interactions on learning. The first part of this section highlights the generality of the basic properties of learning reviewed in first sections. There are many situations in which social behavior can be accurately predicted based on simple models that were designed to capture behavior in individual choice tasks. Yet there are also interesting exceptions to this generality. The main exceptions can be summarized with the assertion that in certain settings prior information changes the strategies that are considered in the learning process.

The chapter concludes with a discussion of the practical implications of experimental learning research. The discussion focuses on the economics of small decisions.

1 THE BASIC PROPERTIES OF DECISIONS FROM EXPERIENCE

The current section reviews the learning phenomena that we consider to be most basic in the sense that they can be reliably replicated in the most basic versions of the clicking paradigm.

1.1 Six Basic Regularities and a Model

Recall that the current review is based on the distinction between three cognitive factors that drive decisions from experience: the cognitive strategies, the exploration policy, and the choice rule. The present subsection tries to clarify the basic properties of the choice rule. In order to achieve this goal, it focuses on phenomena that can be replicated even when the role of sophisticated cognitive strategies and of the exploration policy is minimized. This "minimization" is achieved by using the 2-alternative clicking paradigm with complete feedback (cf. Figure 10.1) and a static payoff rule. After each choice in the clicking experiments considered here, the agents receive feedback concerning their obtained payoff (the payoff from the key selected), and the forgone

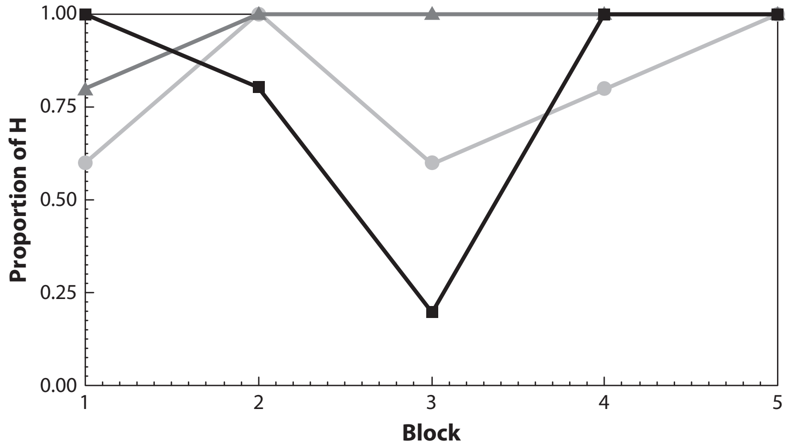

Figure 10.2: The proportion of H choices of three participants in the first 5 blocks of 5 trials of an experiment that involves a choice between a key that provides 1 with certainty (option H) and a key that provides 0 with certainty. All subjects learn to maximize, but the process is stochastic.

payoff (the payoff that could have been obtained had the second key been selected). The payoff of each key is drawn from a payoff distribution associated with that key. For example, if the key is associated with payoff distribution "11 with probability .5, −9 otherwise", the payoff will be 11 in 50% of the trials and −9 in the other 50%. The fact that the payoff rule is static implies that the distributions do not change during the experiment, and the agents can maximize expected return by selecting the option that has led to higher average payoff in the previous trials.

Our review uncovers six robust behavioral regularities that emerge in this setting. All six regularities imply deviations from maximization of expected return. Yet we believe that they do not imply deviations from "ecologically reasonable" behavior. In order to clarify this assertion, we conclude the presentation of each behavioral regularity with an "ecological justification." Section 1.1.7 presents a model that quantifies these justifications, and the subsequent sections clarify the model's relationship to other models and its predictive value.

1.1.1 The Law of Effect

Thorndike (1898) studied how cats learn to escape from puzzle boxes. The experiments included several trials: Each trial started with the placement of a cat in a puzzle box and ended when the cat exited the box. Evaluation of the learning curves (time to escape as a function of trial number) led Thorndike to conclude that the learning was gradual and stochastic. There was no evidence of sudden jumps in performance. Thorndike summarized this observation with the law of effect: Choices that have led to good outcomes in the past are more likely to be repeated in the future.

Studies that use the clicking paradigm reveal a similar pattern. Subjects tend to select the alternative that led to good outcome in the past, and the learning curves appear to reflect a gradual and stochastic process. Figure 10.2 demonstrates this pattern. Each curve in this figure summarizes the behavior of 1 participant in the first 25 trials of a simple experiment. The experiment involved a trivial choice task: one option, referred to as H (high payoff) always provided a payoff of 1 shekel, and the second option always led to a payoff of 0. The experiment used the basic clicking paradigm. That is, the participants did not receive prior information concerning the payoff rule and could rely on feedback concerning the obtained and forgone payoffs. The results, presented in 5 blocks of 5 trials each, reveal that by the last block, all 3 subjects learned to prefer the better option (H). Yet, the learning process is noisy. For example, the proportion of optimal choices of the "circle" subject go up to 100% by the second block, then go down to 60% in the third block, and then go up to 100% in the fifth block.

An ecological justification: Exploration. Recall that the current analysis focuses on conditions that minimize the value of exploration. The agents' actions did not affect their feedback. However, the subjects could not know with certainty that this is the case. Thus, the observed deviations from "best reply to past experience" can be an indication of exploring the effect of selecting the 0 key.

1.1.2 The Payoff Variability Effect

Myers and Sadler (1960) studied decisions from experience using a "card-flipping" paradigm. In each trial of their studies, the participant saw one side of a card and had to decide whether to accept the payoff written on that side (the safe alternative), or the payoff written on the unobserved side of the card (the riskier option). Participants received feedback concerning their payoffs after each choice (the card was flipped only if the participant chose the riskier option). The results revealed that an increase in the payoff variability of the risky option (the variability of the payoff distribution on the unobserved side) reduced the proportion of choices that maximized expected payoff. Busemeyer and Townsend (1993) termed this pattern the "payoff variability effect" and highlighted its robustness.

We replicated this pattern in the clicking paradigm with the study of Problems 1, 2, and 3 (the H-rate in the brackets on the right are the proportion of H choices over all trials, and EV is the expected value of the gamble):

Problem 1 ($r = 200$, $n = 20$, FB = complete, payoff in shekels in a randomly selected trial)

H	1 with certainty	[H-rate: 96%]
L	0 with certainty	

Problem 2 (same procedure as in Problem 1)

H	+11 with probability 0.5 −9 otherwise (EV = 1)	[H-rate: 58%]
L	0 points with certainty	

Problem 3 (same procedure as in Problem 1)

H	0 with certainty	[H-rate: 53%]
L	9 with probability 0.5 −11 otherwise (EV = −1)	

Problems 1, 2, and 3 were run in the same experiment using a within-participant design. Each of 20 participants ($n = 20$) faced each problem for 200 rounds ($r = 200$) under the clicking paradigm with complete feedback (FB = complete). The order of the 3

Problem	H (high EV)	L
1	1 with certainty	0 with certainty
2	(11, 0.5; −9)	0 with certainty
3	0 with certainty	(9, 0.5; −11)

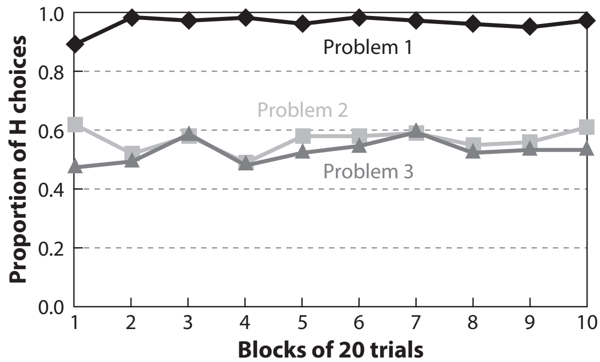

Figure 10.3: Proportion of H choices in problems 1–3 in 10 blocks of 20 trials. The results demonstrate the payoff-variability effect.

problems was random. The participants did not receive a description of the problems but were informed that the experiment includes 3 independent parts and when a new part starts. The final payoff for the experiment was the sum of a show-up fee of 30 shekels and the outcome of one randomly selected trial.

Notice that problems 1 and 2 involve a choice between alternative H, with an EV of 1 shekel, and alternative L, with an EV of 0. The higher EV maximization rate (H-rate) in problem 1 (96%) compared to problem 2 (58%) is suggestive of risk aversion and/or loss aversion (oversensitivity to losses): H was less attractive (in Problem 2) when it increased the variance and was associated with losses. However, the risk-aversion and the loss-aversion explanations are inconsistent with a comparison of problem 2 and problem 3. In problem 3, risk aversion and loss aversion imply maximization (H choices). The results show an H-rate of only 53%. Figure 10.3 presents the observed choice rate of H in blocks of 20 trials. It shows that the differences between the three conditions are relatively robust over time.

Additional studies, reviewed in Erev and Barron (2005), demonstrate the robustness of the payoff variability effect. These studies reveal robustness to the payoff sign, to incomplete feedback, and to the number of possible outcomes.[5]

Chasing, the big eyes effect, and contingent loss aversion. One reasonable explanation of the results in problems 1–3 involves the assertion of large individual differences in risk attitude and/or in the attitude toward losses. For example, the aggregate results are consistent with the hypothesis that about half the participants are risk averse and the other half are risk seekers. However, this explanation has important shortcomings. One clear shortcoming is the fact that the correlation between the R-rate in problems 2 and 3 is not large (see Section 1.1.6). A more interesting shortcoming is suggested by studies that examine investment decisions. These studies show that investors tend to "chase" past returns (see Kliger, Levy, and Sonsino 2003; Grinblatt, Titman, and Wermers 1995). That is, they tend to invest in assets that led to high earnings in the past. Grosskopf, Erev, and Yechiam (2006) show that this "big eyes effect" implies that payoff variability can lead most agents to behave as if they are risk seekers. Ben Zion et al. (2010) clarify the robustness of this observation in a study that focuses on the following problem:

A simplified investment problem ($r = 100$, $n = 30$, FB = complete, 1 point = 0.25¢, pay rule = random trial)

R1	$4x$ (EV = 0)	
R2	$2y - 2x$ (EV = 0)	
S	$x + y + 5$ (the mean of R1 and R2 plus 5, EV = 5)	[S-rate = 25%]

Here, x is a draw from a normal distribution with a mean of 0 and standard deviation of 300 ($x \sim N(0, 300)$), and y is a draw from a normal distribution with a mean of 0 and standard deviation of 10 ($y \sim N(0,10)$).

Ben Zion et al. study can be described as a simulation of a simplified investment task. Options R1 and R2 simulate two risky stocks, and option S simulates an attractive index fund that provides the mean of R1 and R2, plus a small bonus. Thus, option S has the highest mean and lowest variance. The experiment used the clicking paradigm with complete feedback. In addition, the participants received a complete description of the payoff rule. The description emphasized the fact that S provides the mean of R1 and R2 plus 5.

The results reveal random choice in the first trial (S-rate of 33%), and a decrease in the tendency to select S with experience. That is, experience with the high payoff variability investment problem impaired maximization. The S-rate in the last block of 20 trials was only 18%. This value is much lower than the 50% rate implied by the assertion that about half of the participants are risk and/or loss averse and lower than the 33% implied under random choice.

The correlation effect. Diederich and Busemeyer (1999) highlight an important boundary condition for the payoff variability effect. When the payoffs of the different alternatives are positively correlated, the availability of information concerning foregone payoffs eliminates the payoff variability effect. In the extreme case in which alternative H dominates L in all trials, payoff variability has little effect.

Grosskopf, Erev, and Yechiam (2006) demonstrate the robustness of this "correlation effect" in the clicking paradigm. They focused on the following two problems:

Problem 4 ($r = 200$, $n = 10$, FB = complete, accumulated payoffs 10 units = 0.01 shekel)

H	$N(120, 10) + c_t$(EV = 120)	[H-rate: 75%]
L	$N(100, 10) + d_t$ (EV = 100)	

Problem 5 (same procedure as in Problem 4)

H	$N(120, 10) + c_t$ (EV = 120)	[H-rate = 98%]
L	$N(100, 10) + c_t$ (EV = 100)	

The exact payoffs were the rounded sum of two terms: A draw from a normal distribution with a mean of 100 or 120 and standard deviation of 10, and (c_t or d_t), a draw from the distribution (-50 with $p = \frac{1}{3}$; 0 with $p = \frac{1}{3}$; $+50$ otherwise). The values of c_t and d_t were independent. Thus, the payoffs of the two alternatives are positively correlated in problem 5 but are not correlated in problem 4. The feedback after each trial was complete: The participants saw the obtained and the forgone payoffs. The final payoff was the sum of the obtained payoffs, with the conversion rate of 1 shekel per 1,000 points. The results show a clear correlation effect. The correlation increased the maximization rate from 75% (in problem 4) to 98% (in problem 5). Thus, when the correlation is high, subjects can learn to maximize expected return.

Probability learning, matching, and overmatching. Many of the early studies of decisions from experience used the probability-learning paradigm. In each trial of a typical study, the participants are asked to guess if a target lightbulb will flash. The probability of a flash is kept constant throughout the experiment. Correct predictions lead to a small gain, and incorrect predictions lead to a lower payoff (0 or a small loss). Grant, Hake, and Hornseth (1951) found an almost perfect match between the true flash probability and the probability of the choice of yes in trials 55 to 60 of their "training phase." For example, when the probability of a flash was 0.75, the proportion of yes choices in the last block was 75%. Notice that this behavior reflects deviation from maximization: when the probability of flash is 0.75, maximizing reinforcement requires 100% yes choices.

This deviation from maximization, known as *probability matching*, triggered influential studies and lively debates (see Estes 1950, 1964; Bush and Mosteller 1955; Suppes and Atkinson 1960; Edwards 1961; Siegel and Goldstein 1959; Lee 1971; and recent analysis in Bereby-Meyer and Erev 1998; Vulkan 2000; Shanks, Tunney, and McCarthy 2002). The accumulated results demonstrate that probability matching is not a steady state. That is, longer experience slowly moves choice toward maximization. It seems that behavior reflects overmatching: it falls between probability matching and maximization. In animal studies as well (e.g., Sutherland and Mackintosh 1971; Kagel Battalio, and Green 1995), the frequency with which the better alternative is chosen usually exceeds the probability of reinforcement of that alternative. These results imply that behavior in probability learning tasks can be described as an example of the payoff variability effect: when the payoff variability is large, learning is slow and the decision makers do not learn to maximize expected return.

A demonstration of the common findings using the basic clicking paradigm is provided with the study (Ert and Bereby-Meyer, forthcoming) of the following problem:

Problem 6 (r = 500, n = 20, FB = complete, accumulated payoffs, 1 unit = 0.01 shekel)

H	4 if event E occurs; 0 otherwise (EV = 2.8)	[H-rate: 90%]
L	4 if event E does not occur; 0 otherwise (EV = 1.2)	

Here $P(\text{E}) = 0.7$ The observed H-rate was 70% in the first 50 trials, around 90% between trials 51 and 150, and 93% between trial 401 and trial 500.

An ecological justification: Reliance on small samples and similarity based reasoning. The payoff variability and correlation effects can be captured with the assertion that the subjects tend to rely on small samples of past experiences (see Erev and Barron (2005) and related observations in Fiedler (2000); Kareev (2000); Osborne and Rubinstein (1998)). For example, a subject that relies on a sample of 4 observations in trial t, recalls 4 past trials and selects the option that led to the best mean payoff in these trials. The expected H-rate (proportion of H choices) of this hypothetical subject is 100% in problem 1, 69% in problems 2 and 3, 74.6% in problem 4, 99.8% in problem 5, and 90% in problem 6.

Reliance on small samples is ecologically reasonable under two common conditions. First, in many settings, reliance on small samples saves cognitive efforts (Fiedler 2000; Hertwig and Pleskac 2010; Kareev 2000). It is easier to recall small samples, and it is easier to reach clear conclusions. This cognitive benefit is particularly clear when people rely on the small set of the most recent past experiences. A second set of common conditions in which reliance on small samples is reasonable involves situations in which the payoff rule depends on the state of the word, and the world can be in one of many states. The optimal strategy in this setting requires a focus on past experiences that were obtained under the current state and giving less attention to other past experiences. Whereas this strategy requires a rich memory based on complex computations, people appear to follow it (see Gonzalez, Lerch, and Lebiere 2003; Plonsky, Teodorescu, and Erev 2015). And when the state of the world does not change (the situations just considered) it can lead to deviations from maximization.

1.1.3 Underweighting of Rare Events and the Experience-Description Gap

Kahneman and Tversky (1979) demonstrate that two of the best-known violations of mainstream economic theory, the tendency to buy both insurance and lotteries (Friedman and Savage 1948) and the Allais paradox (Allais (1953) and the next section), can be explained as indications of overweighting of rare events. Their influential analysis includes two steps: They first replicated the classical violations in a standardized experimental paradigm and then proposed a model (prospect theory) that captures the two phenomena. Prospect theory captures the two phenomena with the assumption of a weighting function that reflects oversensitivity to rare events (events whose probability is below 0.25).

The standardized paradigm used by Kahneman and Tversky focuses on "decisions from description": the subjects were presented with a precise description of two prospects and were asked to select (once, and without any feedback) the prospect they prefer. Barron and Erev (2003) have examined if these phenomena also emerge in the clicking paradigm. Their original hypothesis was that experience will reduce the magnitude of the deviations from maximization. The results surprised them: In several of the problems that they examined, experience did not enhance maximization. In some cases, experience led to a reversal of the deviations captured by prospect theory: It triggered underweighting of rare events. This pattern is known as the experience-description gap (see the review in Hertwig and Erev 2009).

Problems 7 and 8 demonstrate the evidence for underweighting of rare events in decisions from experience. These problems were studied by Nevo and Erev (2012) using the clicking paradigm with complete feedback. The participants were paid (in shekels) for one randomly selected trial:

Problem 7 ($r = 100$, $n = 48$, FB = complete, payoff in shekels in a randomly selected trial)

S	0 with certainty	[S-rate = 43%]
R	+1 with probability 0.9; −10 otherwise (EV = −0.1)	

Problem 8 (same procedure as in Problem 7)

S	0 with certainty	[S-rate = 72%]
R	+10 with probability 0.1; −1 otherwise (EV = +0.1)	

Notice that in problem 7, the safer option has higher expected value, but the participants tend to select the gamble. Problem 8 reflects the opposite risk preference: The gamble has higher expected value, but the participants tend to select the safer option. As noted by Barron and Erev, this pattern can be a reflection of insufficient sensitivity to the rare and extreme outcomes that occur in 10% of the trials. Thus, the participants behave as if they believe that "it won't happen to me."

The reversed certainty effect (reversed Allais paradox). A clear demonstration of the significance of the difference between decisions from experience and decisions from description is provided by the study of variants of Allais' (1953) common ratio problems. Expected utility theory (von Neumann and Morgenstern 1947) implies that if prospect B is preferred to A, then any probability mixture (B, p) must be preferred to the mixture (A, p).[6] In his classic experiment, Allais (1953) found a clear violation of this prediction. He constructed an example in which the more risky of two prospects becomes relatively more attractive when the probability of winning in both prospects is transformed by a common ratio. Kahneman and Tversky (1979) refer to this pattern as the "certainty effect." Barron and Erev (2003) demonstrate that decisions from experience (in the clicking paradigm with incomplete feedback) reflect the opposite pattern. The study of problems 9 and 10 replicates these results using the clicking paradigm with complete feedback:

Problem 9 (r = 400, n = 24, FB = complete, accumulated payoff 1 point = 0.01 shekel)

S	3 with certainty	[S-rate = 36%]
R	4 with probability 0.8; 0 otherwise (EV = 3.2)	

Problem 10 (same procedure as in Problem 9)

S	3 with probability 0.25; 0 otherwise (EV = 0.75)	[S-rate = 51%]
R R	4 with probability 0.2; 0 otherwise (EV = 0.80)	

The results reveal a reversed certainty effect. The safe option (S) was less attractive in problem 9—when it was associated with certainty—than in problem 10—when it was not. This pattern is consistent with the assertion that in decisions from experience, the least likely events (probability of 0.2) are underweighted.

Additional studies of the certainty effect reveal differences between rats, bees, and human subjects. MacDonald et al. (1991) show that rats exhibit the original certainty effect: They studied variants of problems 9 and 10 with payoffs in cups of water and found *more* S choices when S provides medium pay with certainty. In contrast, Shafir et al. (2008) show that honeybees exhibit the reversed certainty effect. Their study examined variants of problems 9 and 10 with payoffs in terms of the percentage of sugar water and found *fewer* S choices when S provides medium pay with certainty. Shafir et al. suggested that perceptual noise might be responsible. According to this explanation, the rats (but not the bees) had difficulty in discriminating the medium and high payoffs and for that reason preferred S in the variant of problem 9. The value of this explanation was demonstrated in a study with human subjects, which revealed that a reduction in the clarity of the feedback (in a study of problems 9 and 10) leads to the emergence of the original certainty effect.

Underweighting and overestimation. The suggestion that people underweight rare events appears to be inconsistent with previous research that demonstrates overestimation of rare events (e.g., Viscusi 2002; Erev, Wellsten, and Budescu 1994). For example Viscusi (2002) found that both smokers and nonsmokers tend to overestimate the probability that smokers will develop lung cancer. Barron and Yechiam (2009) examined whether this difference between smokers and nonsmokers is mediated by different settings (e.g., clicking vs. smoking) or different tasks (deciding or estimating). They studied problem 11 using the clicking paradigm with complete feedback and one addition: starting at trial 201, the participants were asked to estimate the probability of the rare outcome (1 point with probability 0.15) before each choice. The results reveal a strong tendency to prefer the risky prospect (R) in all 400 trials (mean R-rate of 79%). This result is consistent with underweighting of rare events. The estimations, on the other hand, reflected oversensitivity to rare events. The average estimate (of the 15% event) was 21%. Thus, participants appear to exhibit oversensitivity to rare events in estimation, and undersensitivity to rare events in choice (similar results are reported by Friedman and Massaro, 1998).

Problem 11 (r = 400, n = 24, FB = complete, accumulated payoffs, 1 unit = 0.01 shekel)

R	3 with probability 0.85; 1 otherwise (EV = 2.7)	[R-rate = 79%]
S	2.7 with certainty	

The sampling paradigm and robustness to the number of repeated gamble realizations. Hertwig et al. (2004) note that the "experience-description gap" just summarized can be attributed to two differences between the experimental paradigms: the source of the information (experience or description), and the number of repeated realizations of the gambles (one or many). To evaluate the role of these factors, they examined some of the problems considered by Barron and Erev (2003) under two conditions: one-shot decisions from description and one-shot decisions from experience.

The two conditions differed only with respect to how the decision makers learned about the options' outcomes and likelihoods. In the description group, options were described as in Kahneman and Tversky's studies.

In the sampling group, the information describing the options was not displayed. Instead, participants were shown two buttons on the computer screen and were told that each button was associated with a payoff distribution. Pressing on a given button elicited the sampling of an outcome (with replacement) from its distribution. In problem 9, for example, drawing from one distribution led to the outcome 4 in 80% of all draws and to the outcome 0 in 20% of all draws. Sampling from the other distribution always resulted in the outcome 3. Participants could sample however often they wished. By repeatedly experiencing the contingency between choices and outcomes, participants could gradually acquire knowledge about the options' payoff structure. Once they stopped sampling, they indicated their preferred option, and, after completing all problems, participants received monetary payoffs according to their choices and the outcomes of the draws.

The observed choice proportions in the sampling group exhibit the pattern observed under the study of the same problems by Barron and Erev (2003) using the clicking paradigm. That is, the participants behave *as if* they underweight rare events. The correlation between the sampling and the clicking results was 0.92. The observed choice proportion in the description group exhibits the pattern predicted by prospect theory—the participants behave as if they overweight rare events. The correlation between the sampling and the description group was −0.67. These results (and similar findings reported in Weber, Shafir, and Blais (2004), Ungemach , Chater, and Stewart (2008). Erev, Ert, et al. (2010a), Hau, Pleskac, and Hertwig (2008) and in reviews by Hertwig and Erev (2009) and Rakow and Newell (2010)) suggest that the tendency to underweight rare events can be observed in one-shot decisions from experience. Thus, the distinct information source is a sufficient condition for the experience-description gap.

Robustness to prior information. Lejarraga and Gonzalez (2011) have examined the effect of prior information concerning payoff distributions on the tendency to underweight rare events in the clicking paradigm, examining the joint effect of description and experience. In one of their studies, the participants were asked to select between a safe prospect that provides 3 with certainty and a gamble that provides 64 with probability 0.05 and 0 otherwise. Their results reveal that the initial behavior reflects high sensitivity to the rare events, with the emergence of underweighting of rare events with experience. The proportion of gambles chosen between trial 10 and 100 was below 30%. Jessup, Bishara, and Busemeyer (2008) document a similar pattern in a study in which the exact value of the gamble varied from trial to trial. Alternative explanations of the weak effect of description of the incentive structure, in the current setting are discussed in Section 1.1.9.

Sensitivity to expected values. An extreme interpretation of the results just summarized would be that decision makers tend to neglect rare events; that is, in most cases they fail to consider these events. Ert and Erev (2016) show a shortcoming of this extreme explanation by examining the following problems:

Problem 12 (r = 400, n = 24, FB = complete, accumulated payoffs, 1 unit = 0.01 shekel)

H	2.52 with certainty	[H-rate = 40%]
L	2.53 with probability 0.89; 2.43 otherwise (EV = 2.519)	

Problem 13 (same procedure as in Problem 12)

H	2.52 with certainty	[H-rate = 72%]
L	2.53 with probability 0.89; 2.03 otherwise (EV = 2.48)	

The results show a deviation from maximization consistent with underweighting of rare events in problem 12 but not in problem 13. This pattern suggests that the rare events are not neglected. When they are sufficiently important they are taken into account.[7]

Sensitivity to the coefficient of variance. Shafir (2000) reviews experimental studies of animal risk attitude in a binary choice task. The results suggest that under normal conditions the tendency to select the safer alternative is better predicted by the coefficient of variance (CV) than by the variance of the risky alternative. CV is defined as the payoff standard deviation divided by the payoff mean. Weber, Shafir, and Blais (2004) show that this pattern is consistent with underweighting of rare events. Underweighting of rare events implies risky choices when the CV is low (relatively high mean) and risk aversion when the CV is high (relatively low mean).

Signal-detection tasks. In binary signal–detection tasks, an observer is asked to classify stimuli that belong to one of two distributions. In a typical experiment (see the review in Erev 1998), the two distributions are normal with equal variance, and they represent the state of the world. For example, the state may be the gender of a candidate (male or female), and the signal may be the candidate's height. After each response (guessing male or female) the observer receives immediate payoff determined by a fixed 2×2 payoff matrix that gives the payoff for each of the four possible outcomes (correct detection of a male, correct detection of a female, incorrect male response, and incorrect female response). Assuming that the male's mean is higher, the optimal choice rule is a cutoff strategy of the type respond male if the signal exceeds a certain height. The location of the cutoff depends on the payoff of the 4 outcomes and on the prior probability of the two distributions. Experimental studies of this task reveal higher sensitivity to the prior probabilities than to the payoffs (see Healy and Kubovy 1981). Barkan, Zohar, and Erev (1998) show that this pattern implies deviation from maximization in the direction of underweighting rare events.

An ecological justification: Sampling and weighting. The tendency to underweight rare events can be explained with the assertion, just presented, that people rely on small samples of past experiences. For example, a subject that relies on a sample of 4 past experiences will prefer the negative EV gamble -10 with probability 0.1, $+1$ otherwise over 0 with certainty in 56% of the trials (because 65% of the samples of size 4 do not include the 10% event). However, this assumption of strong reliance on small samples cannot explain the observed sensitivity to the expected value in problem 13 (i.e., reliance on a sample of 4 implies the same behavior in problems 12 and 13). The coexistence of underweighting of rare events and some sensitivity to expected values can be captured by a weak variant of the reliance on the small-samples hypothesis: the assumption that a small sample of experiences receives more attention than the other experiences, but all experiences receive some attention. Thus, when the difference in the expected values is large enough, it affects behavior.

1.1.4 The Very Recent and the Wavy Recency Effects

Analysis of the effect of recent outcomes on choice behavior in probability learning tasks led Estes (1964; also see the review in Lee (1971)) to conclude that the most common pattern is positive recency: decision makers are more likely to select the alternative that led to the best outcome in recent trials.

A clear example of positive recency in the clicking paradigm is provided in the analysis of the contingent choice rate in problems 2 and 3 in the top panel of Table 10.1. The probability of risky (R) choices is larger, in these problems, after high payoff from R than after low payoff from R. The overall R-rates are 64% after high payoff and 40% after low payoff. Aggregation over the two payoffs (high and low) suggests that that the proportion of choices that are the best reply to the most recent payoff, referred to as Best-Reply-1, is 62%.

TABLE 10.1:
Summary of experiments that examine a choice between a safe prospect and a prospect with no more than two outcomes using the basic clicking paradigm. The recency effects (in bold) are estimated as the difference between the R-rates after high and low payoffs from R given the same recent choice.

			Experimental Results					*The Predictions of I-SAW*				
			R-Rates and Implied Recency Effect as a Function of Last Choice and Recent Payoff from R				*R-Rate over All Trials*	*R-Rates and Implied Recency Effect as a Function of Last Choice and Recent Payoff from R*				
								Last Choice				
		Last choice	S *(R-rate is switch rate, recent payoff from R is forgone)*		R *(R-rate is repetition rate, recent payoff from R is obtained)*			S *(R-rate is switch rate, recent payoff from R is forgone)*		R *(R-rate is repetition rate, recent payoff from R is obtained)*		*R-Rate over All Trials*
	Problem [Number of trials]	*Most recent payoff from R*	*High*	*Low*	*High*	*Low*		*High*	*Low*	*High*	*Low*	
1	S 0 with certainty	R-rate	0.43	—	0.99	—	0.96	0.48	—	0.94	—	0.89
[200]	R 1 with certainty		—		—			—		—		
2	S 0 with certainty	R-rate	0.56	0.21	0.81	0.59	0.58	0.37	0.30	0.81	0.76	0.61
[200]	R 1 (11, 0.5; −0.9)	**Recency**	**+0.35**		**+0.22**			**+0.07**		**+0.05**		
3	S 0 with certainty	R-rate	0.40	0.16	0.77	0.60	0.47	0.24	0.18	0.72	0.64	0.40
[200]	R (9, 0.5; −11)	**Recency**	**+0.24**		**+0.17**			**+0.06**		**+0.08**		
4	S 0 with certainty	R-rate	0.23	0.06	0.60	0.79	0.29	0.32	0.13	0.74	0.76	0.38
[100]	R (10, 0.1; −1)	**Recency**	**+0.17**		**−0.19**			**+0.19**		**−0.02**		
8	S 0 with certainty	R-rate	0.21	0.31	0.84	0.69	0.56	0.24	0.26	0.87	0.68	0.62
[100]	R (1, 0.9; −10)	**Recency**	**−0.10**		**+0.15**			**−0.02**		**+0.19**		
9	S 3 with certainty	R-rate	0.2	0.20	0.91	0.67	0.64	0.36	0.38	0.85	0.76	0.68
[400]	R (4, 0.8; 0)	**Recency**	**+0.06**		**+0.24**			**−0.02**		**+0.09**		

TABLE 10.1:
Continued.

			Experimental Results: R-Rates and Implied Recency Effect as a Function of Last Choice and Recent Payoff from R				*Experimental Results: R-Rate over All Trials*	*The Predictions of I-SAW: R-Rates and Implied Recency Effect as a Function of Last Choice and Recent Payoff from R*				*The Predictions of I-SAW: R-Rate over All Trials*
		Last choice	S *(R-rate is switch rate, recent payoff from R is forgone)*		R *(R-rate is repetition rate, recent payoff from R is obtained)*			Last Choice: S *(R-rate is switch rate, recent payoff from R is forgone)*		Last Choice: R *(R-rate is repetition rate, recent payoff from R is obtained)*		
	Problem [Number of trials]	*Most recent payoff from R*	*High*	*Low*	*High*	*Low*		*High*	*Low*	*High*	*Low*	
12	S 2.52 with certainty	R-rate	0.15	0.09	0.94	0.78	0.60	0.28	0.30	0.85	0.72	0.63
[400]	R (2.53, 0.89; 2.43)	**Recency**	**+0.06**		**+0.16**			**−0.02**		**+0.13**		
13	S 2.52 with certainty	R-rate	0.06	0.08	0.92	0.63	0.28	0.08	0.08	0.76	0.57	0.24
[400]	R (2.53, 0.89; 2.03)	**Recency**	**−0.02**		**+0.29**			**0**		**+0.19**		
14	S 7 with certainty	R-rate	0.40	0.04	0.94	0.95	0.45	0.46	0.07	0.77	0.91	0.46
[100]	R (16.5, 0.01; 6.9)	**Recency**	**+0.36**		**−0.01**			**+0.39**		**−0.14**		
15	S −9.4 with certainty	R-rate	0.15	0.06	0.70	0.80	0.26	0.23	0.09	0.56	0.71	0.26
[100]	R (−2, 0.05; −10.4)	**Recency**	**+0.09**		**−0.10**			**+0.14**		**−0.15**		
16	S −4.1 with certainty	R-rate	0.27	0.06	0.86	0.94	0.54	0.36	0.11	0.81	0.84	0.42
[100]	R (1.3, 0.05; −4.3)	**Recency**	**+0.21**		**−0.08**			**+0.25**		**−0.03**		
17	S −18.7 with certainty	R-rate	0.29	0.06	0.85	0.87	0.38	0.31	0.11	0.72	0.76	0.35
[100]	R (−7.1, 0.07; −19.6)	**Recency**	**+0.23**		**−0.02**			**+0.20**		**−0.04**		
18	S −7.9 with certainty	R-rate	0.20	0.06	0.86	0.84	0.31	0.30	0.12	0.71	0.75	0.34
[100]	R (5, 0.08; −9.1)	**Recency**	**+0.14**		**+0.02**			**+0.18**		**−0.04**		

TABLE 10.1:
Continued.

			Experimental Results					*The Predictions of I-SAW*				
			R-Rates and Implied Recency Effect as a Function of Last Choice and Recent Payoff from R				*R-Rate over All Trials*	*R-Rates and Implied Recency Effect as a Function of Last Choice and Recent Payoff from R*				
								Last Choice				
		Last choice	S *(R-rate is switch rate, recent payoff from R is forgone)*		R *(R-rate is repetition rate, recent payoff from R is obtained)*			S *(R-rate is switch rate, recent payoff from R is forgone)*		R *(R-rate is repetition rate, recent payoff from R is obtained)*		*R-Rate over All Trials*
	Problem [Number of trials]	*Most recent payoff from R*	*High*	*Low*	*High*	*Low*		*High*	*Low*	*High*	*Low*	
19	S −25.4 with certainty	R-rate	0.22	0.07	0.89	0.90	0.45	0.35	0.13	0.82	0.83	0.46
[100]	R (−8.9, 0.08; −26.3)	**Recency**	**+0.15**		**−0.01**			**+0.22**		**−0.01**		
20	S 11.5 with certainty	R-rate	0.29	0.07	0.81	0.78	0.30	0.23	0.11	0.60	0.67	0.26
[100]	R (25.7, 0.1; 8.1)	**Recency**	**+0.22**		**+0.03**			**+0.12**		**−0.07**		
21	S −15.5 with certainty	R-rate	0.42	0.19	0.91	0.75	0.68	0.44	0.39	0.90	0.86	0.77
[100]	R (−8.8, 0.6; −19.5)	**Recency**	**+0.23**		**+0.16**			**+0.05**		**+0.04**		
22	S 2.2 with certainty	R-rate	0.13	0.15	0.85	0.68	0.47	0.25	0.30	0.89	0.71	0.67
[100]	R (3, 0.93; −7.2)	**Recency**	**−0.02**		**+0.17**			**−0.05**		**+0.18**		
23	S 25.2 with certainty	R-rate	0.14	0.32	0.86	0.82	0.52	0.25	0.31	0.90	0.71	0.68
[100]	R (26.5, 0.94; 8.3)	**Recency**	**−0.18**		**+0.04**			**−0.06**		**+0.19**		
24	S 6.8 with certainty	R-rate	0.08	0.23	0.92	0.77	0.50	0.16	0.21	0.90	0.65	0.60
[100]	R (7.3, 0.96; −8.5)	**Recency**	**−0.15**		**+0.15**			**−0.05**		**+0.25**		
25	S 11 with certainty	R-rate	0.09	0.19	0.94	0.71	0.57	0.19	0.33	0.92	0.70	0.68
[100]	R (11.4, 0.97; 1.9)	**Recency**	**−0.10**		**+0.23**			**−0.14**		**+0.22**		

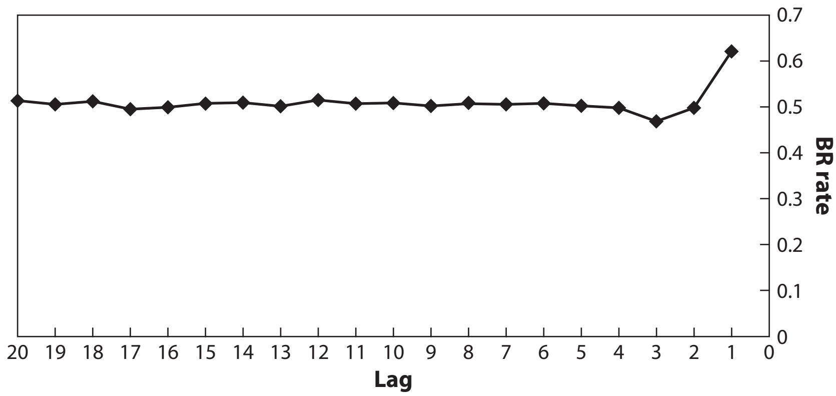

Figure 10.4: The very recent effect in problems 2 and 3: the proportion of choices (at trial t) of the alternative that led to the best outcome in trial t-Lag. Thus, Lag = 1 (*on the right*) presents the best reply rate to the most recent trial.

An extension of this analysis to other recent outcomes reveals an interesting pattern. To describe this pattern, let Best-Reply-L be the choice rate of the alternative that led to the best outcomes exactly L trials before the current trial. Figure 10.4 presents the values of Best-Reply-1 to Best-Reply-20 (based on data from trials 21 up to 200 in problems 2 and 3). The results reveal a large qualitative difference between Best-Reply-1 and the other values. The decrease in the effect of recent outcomes appears to be sharp. Best-Reply-1 reflects a strong recency effect, but Best-Reply-2 and -3 are not larger than the mean value. Indeed, Best-Reply-3 is the lowest point in the curve in Figure 10.4. Nevo and Erev (2012) refer to this pattern as the very recent effect. Plonsky, Teodorescu, and Erev (2015) show that deviation from positive recency is even larger in problems with rare events: the recency curve is problems of this type tend to be wavy.

An ecological justification: State inertia. The unique effect of the most recent outcome can be captured with the assertion that in some trials the decision makers behave as if they assume that the payoff rule is determined by the state of nature, and the current state is not likely to change (the state in the next trial is likely to be identical to the state in the last trial).

1.1.5 Inertia and Surprise-Triggers-Change

Analysis of the relationship between recent and current choice reveals strong positive correlation that implies inertia (Nevin 1988; Cooper and Kagel 2008; Suppes and Atkinson 1960, Erev and Haruvy 2005). Decision makers tend to repeat their last choice. For example, over problems 2 and 3, the participants repeated their last choice in 68% of the trials. Moreover, inertia is a better predictor of behavior than positive recency. When inertia and positive recency lead to contradicting predictions, the decision makers are more likely to exhibit inertia (as noted in Section 1.1.4, the positive recency rate is only 62%).

Overalternation. Previous research highlights two boundary conditions for inertia. First, in some cases human decision makers exhibit overalternation when they are asked to select between alternatives that are known to be identical (see Rapoport and Budescu (1997) and Section 4.2.2). Second, animal studies (see review in Dember and Fowler

1958) highlight spontaneous alternation by certain species in certain settings that can be described as a response to an environment in which inertia is counterproductive.

Negative recency. The first row in Table 10.1 presents the choice rates in problems 7 and 8 by the last choice and the recent payoffs. The results reveal two deviations from positive recency. The first deviation emerges in problem 8 after choice R. The rate of repeated R choice was 79% after a loss (the payoff −1), and only 61% after a gain (payoff of +10). The second indication of negative recency is observed in problem 7 after choice S. The rate of a switch to R was 31% after a forgone loss (the payoff −10), and only 21% after a forgone gain (payoff of +1).

The lower rows in Table 10.1 demonstrate that this pattern is not unique to problems 7 and 8. It presents the results obtained in the study of 12 additional problems by Nevo and Erev (using the basic clicking paradigm). Most problems reveal higher change rates after surprising outcomes, even when the surprising outcomes reinforce the last choice.

The relative effect of obtained and foregone outcomes. Under an extreme interpretation of Thorndike's (1898) law of effect, behavior is driven by obtained outcomes. Thus, information concerning foregone payoffs is not likely to have a significant effect. However, experimental evaluations of this hypothesis show that it can be rejected (e.g., Mookherjee and Sopher 1994, 1997; Camerer and Ho 1999; Nyarko and Schotter 2002; Marchiori and Warglien 2008). In fact, in certain settings people are more sensitive to foregone than to obtained outcomes (e.g., Grosskopf, Erev, and Yechiam 2006). The results, presented in Table 10.1, reveal a similar pattern: the best reply rate to the forgone payoff is larger than the best reply rate to the obtained payoff. One boundary condition to the current observation involves the number of alternatives. When the number of alternatives is very large, people are more likely to pay attention to the payoff of the alternative they chose than to the forgone payoff from each of the other multiple alternatives (see Ert and Erev 2007).

An ecological justification: Action inertia and surprise-trigger-change. The observed inertia and the complex recency pattern documented in Table 10.1 can be captured with the hypothesis that in certain trials people choose an inertia mode and simply repeat their last choice. This tendency is ecologically reasonable when the cost of deciding is larger than the expected benefit. Specifically, if the agent carefully reached a decision before trial t, making another costly decision at trial t is likely to be cost effective only if the recent feedback is surprising.

1.1.6 Individual Differences and the Iowa Gambling Task

While studying patients with neuropsychological disorders, Bechara et al. (1994) have found that a specific neurological syndrome is associated with poor performance in a simple decision-from-experience task. The population they studied involved patients with lesions in the orbitofrontal cortex. This syndrome involves intact IQ and reasoning skills but poor decision-making capacities. The task they proposed for assessing decision capacities is now known as the Iowa gambling task. It is presented as a choice between four decks of cards. Each alternative results in one of two outcomes: a sure gain and some probability of a loss. The implied payoff distributions (the sum of the two outcomes) are as follows.

The Iowa gambling task:
Dis. R: Win \$100 with probability 0.9; lose \$1150 otherwise (EV = −25)
Dis. S: Win \$100 with probability 0.5; lose \$150 otherwise (EV = −25)

Adv. R: Win \$50 with probability 0.9; lose \$200 otherwise (EV = +25)
Adv. S: Win \$50 with probability 0.5; 0 otherwise (EV = +25)

As in the clicking paradigm, the decision makers do not receive a description of the different distributions. Their information is limited to the obtained payoff after each trial. The experiment included 100 trials.

Notice that two of the alternatives are advantageous (Adv. R and Adv. S have expected payoff of 25), and two are disadvantageous (Dis. R and Dis. S have expected value of −25). Bechara et al. found that the patients with lesions in the orbitofrontal cortex did not learn to avoid the disadvantageous alternatives, while the participants in the control groups (patients with other neurological problems) did.

Following up on these findings, Busemeyer and Stout (2002) presented a simple reinforcement learning model that implies that the failure to learn in the Iowa gambling task can be a product of three different behavioral tendencies: overexploration, a recency effect, and insufficient sensitivity to losses. Under Busemeyer and Stout's model, these three tendencies are abstracted as parameters that can be estimated from the data.

Yechiam et al. (2005; 2008) showed the value of this approach. For example, they showed that the estimation of the learning parameters can be used to distinguish between criminals. In their study of first-time offenders at the reception and classification facility for the State of Iowa Department of Corrections, diverse criminal subgroups all performed poorly in the Iowa gambling task. However, it was found that criminals incarcerated for drug addiction or repeat sex offenders showed insufficient sensitivity to losses. In contrast, more-violent criminals, including those convicted of assault and/or murder—and to some extent those convicted of robbery as well— exhibited high recency.

An additional indication of the significance of individual differences is provided by the analysis of the correlation between behavior in problems 2 and 3 in the clicking experiment described earlier. Recall that the experiment used the basic clicking paradigm, and 20 participants faced both problems. Following Yechiam et al. (2005), we focused on three variables: the proportion of risky choices (a measure of attitude toward losses), the proportion of Best-Reply-1 (a measure of a recency effect), and the distance between the mean choice rate and 0.5 (a measure of decisiveness). The observed correlations are 0.18, 0.75, and 0.69 for loss attitude, recency, and decisiveness (with the last two values highly significant).

An ecological justification: Variability facilitates evolution and learning. The existence of variability is a necessary condition for survival in a number of instances, so it would be selected in the evolutionary process. Another attractive feature of variability in learning is the fact that it can facilitate coordination. Specifically, variability enhances efficiency in coordination games in which the payoff decreases with the number of people that make the same choice. One example is the market-entry game described later.

1.1.7 Quantitative Summary: Inertia, Sampling, and Weighting (I-SAW)

Nevo and Erev (2012) propose a descriptive model that can reproduce the six behavioral regularities just presented. The model, referred to as I-SAW, can be described by the following assumptions.

I-SAW1: Three response modes. The model distinguishes between three response modes: exploration, exploitation, and inertia. Exploration is assumed to imply random choice. The probability of exploration by individual i is set to 1 in the first trial and ε_i (a trait of i) in all other trials.

During exploitation trials, individual i selects the alternative with the highest estimated subjective value (ESV). The ESV of alternative j in trial $t > 1$ is

$$\mathrm{ESV}(j, t) = (1 - w_i)(\mathrm{S\_Mean}) + w_i(\mathrm{G\_Mean}) \tag{1}$$

where S_Mean (sample mean) is the average payoff from alternative j in a small sample of μ_i previous trials in similar settings, G_Mean (grand mean) is the average payoff from j over all $(t-1)$ previous trials, and μ_i and w_i are traits.

The assumed reliance on a small sample from experience is introduced to capture underweighting of rare events and the payoff variability effect (see similar abstractions in and related ideas in Osborne and Rubinstein 1998; Fiedler 2000; Kareev 2000; Rapoport and Budescu 1997; Hertwig et al. 2004; Lebiere, Gonzalez, and Martin 2007). The assumed sensitivity to the grand mean was introduced (following a similar assumption in Gonzalez et al. 2003) to capture the observed sensitivity to expected values.

I-SAW2: Similarity and recency. The μ_i draws are assumed to be independent (sampling with replacement) and biased toward the most recent experience (trial $t-1$). A bias with respect to trial $t-1$ occurs with probability ρ_i (a trait). The remainder of the time (probability $1-\rho_i$), the agent relies on experience from the trials that appear to be most similar to the current trial. When all the previous trials are objectively equally similar (the current case), the apparent similarity criterion implies random choice. The motivation behind the recency bias is the "very recent effect."

I-SAW3: Surprise-triggers-change. Inertia is added with the assumption that the individuals tend to repeat their last choice. The exact probability of inertia at trial $t+1$ is assumed to decrease when the recent outcomes are surprising. Specifically, if the exploration mode was not selected, the probability of inertia is

$$P(\text{inertia at } t+1) = \pi_i^{\mathrm{surprise}(t)} \tag{2}$$

where $0 \le \pi_i < 1$ is a trait that captures the tendency for inertia. The value of the surprise term is assumed to equal the average of four gaps between certain expectations and the obtained payoffs. In the first two (one for each alternative), the assumed expectation is that the last payoff will be obtained again; thus the gap is between the payoff at $t-1$ and the payoff at t. In the last two, the assumed expectation is the mean payoff; thus, the gap is between the grand mean and the payoff at t. Specifically,

$$\mathrm{gap}(t) = \frac{1}{4}\left[\sum_{j=1}^{2}\left|\mathrm{obtained}_j(t-1) - \mathrm{obtained}_j(t)\right| + \sum_{j=1}^{2}\left|G\_\mathrm{mean}_j(t) - \mathrm{obtained}_j(t)\right|\right] \tag{3}$$

where $\mathrm{obtained}_j(t)$ is the payoff obtained from j at trial t, and $G\_\mathrm{mean}_j(t)$ is the average payoff obtained from j in the first $t-1$ trials (the grand mean). The surprise at t is normalized by the mean gap (in the first $t-1$ trials):

$$\text{surprise}(t) = \frac{\text{gap}(t)}{[\text{mean\_gap}(t) + \text{gap}(t)]} \tag{4}$$

The mean gap at t is a running average of the gap in the previous trials (with mean_gap(1) = 0.00001). Specifically,

$$\text{mean\_gap}(t+1) = \text{mean\_gap}(t)(1 - 1/r) + \text{gap}(t)(1/r) \tag{5}$$

where r is the expected number of trials in the experiment (100 in the current study).

Notice that the normalization (equation 4) implies that the value of surprise(t) is between 0 and 1, and the probability of inertia is between π_i (when surprise$(t) = 1$) and 1 (when surprise $(t) = 0$). An interesting justification for this gap-based abstraction comes from the observation that the activity of certain dopamine-related neurons is correlated with the difference between expected and obtained outcomes (see Schultz (1998) and related analysis in Caplin and Dean (2007)).

I-SAW4: Individual differences, traits, and parameters. The traits are assumed to be independently drawn from a uniform distribution between the minimal possible value (allowed by the model) and a higher point. Thus, the model has 5 free parameters: the highest point of the 5 distributions.

Estimation and results. We used a grid-search procedure to estimate the parameters of the model. The criterion was the mean-squared deviation (MSD) between the model's predictions and the experimental results (including the results summarized in Table 10.1). That is, we ran computer simulations to derive the predictions of the model under different parameters and selected the parameters that minimize the MSD score. The estimated parameters imply the following trait distribution: $\varepsilon_i \sim U[0, 0.24]$, $w_i \sim U[0, 1]$, $\rho_i \sim U[0, 0.12]$, $\pi_i \sim U[0, 0.6]$, and $\mu_i = 1, 2, 3$, or 4.

The right-hand columns in Table 10.1 present the predictions of I-SAW with these parameters. The results reveal that I-SAW reproduces all the behavioral tendencies documented in Table 10.1. In addition, the model provides good quantitative fit. For example, the correlation between the predicted and the observed aggregate choice rates is 0.9, and the MSD score is 0.007. Additional evaluations of this model are discussed in Sections 1.3 and 4.2.

1.1.8 Implications of I-Saw Relative to Traditional Reinforcement Learning and Fictitious Play Models

I-SAW can be described as an example of a reinforcement-learning model (see Satton and Barto 1998; Roth and Erev 1995) and as a generalization of the fictitious play rule (Brown 1951; Robinson 1951; Fudenberg and Levine 1998) and of the naïve sampler model (Erev and Roth 2014). The following section clarifies these connections.

Fictitious play (FP). The FP rule assumes that the decision maker tries to maximize expected return under the assumption that the payoff distributions are static. This assumption is fictitious in many settings, but it is correct in the basic clicking paradigm. At trial t>1, this rule implies a selection of the alternative that led to the highest average payoff in the first $t-1$ trials (and random choice is assumed at $t = 1$). I-SAW implies FP with the traits: $\varepsilon_i = 0$, $w_i = 1$, $\rho_i = 0$, and $\pi_i = 0$. That is, under the FP rule, the estimated subjective value is the grand mean (G_mean), and the alternative with the highest G_mean is selected. The correlation between the aggregate choice rate and the model with these parameters is 0.76, and the MSD score is 0.039. These results suggest that the FP rule (and the implied maximization assumption) provides a useful approximation of the results, but the I-SAW generalization of this rule provides a much better approximation. Additional analysis reveals that the advantage of the generalized model over the FP rule decreases when the difference between the average payoffs from the different alternative is large (relatively to the payoff variability); when this relative difference is large enough the predictions of I-SAW are identical to the predictions of the FP rule.

Stochastic fictitious play (SFP). The SFP model (Cheung and Friedman 1997, 1998) is a generalization of the FP rule that allows for the possibility that the estimated subjective value of option j at trial t includes error. That is,

$$\mathrm{ESV}(j, t) = \mathrm{G\_mean}(j, t) + \varepsilon_{jt} \tag{6}$$

The traditional implementation adds the assumption that the error terms are randomly, identically, and independently distributed. It is convenient to assume that this distribution follows a type I extreme value distribution, which approximates the normal distribution. As demonstrated by McFadden (1974), this assumption implies that the probability of preferring j over k at trial t is

$$P(j, t) = \frac{1}{1 + e^{\sigma[\mathrm{G\_mean}(k,t) - \mathrm{G\_mean}(j,t)]}} \tag{7}$$

SFP can be described as a variant of I-SAW with the parameters $\varepsilon_i = 0$, $w_i = 0.5$, $\rho_i = 0$, and $\pi_i = 0$ and with a modified error term. The error term under I-SAW is determined by a draw of μ_i past experiences. The I-SAW error term is less convenient to model (as it does not allow the derivation of the elegant choice probability term implied under a normal error), but it appears to fit the data better. The advantage of the I-SAW error term is clarified by a comparison of problems 1 and 2. I-SAW implies no error in problem 1 (the trivial no variability problem), and high error rate in problem 2. The SFP allows for the possibility of different error terms by assuming situation specific σ values but cannot predict a long-term difference between the two problems without problem specific parameters.

Reinforcement learning. Simple reinforcement learning models were found to provide good ex ante predictions of behavior in certain games (Erev and Roth 1998), to imply maximization in certain settings (Sutton and Barto 1998), and to be consistent with known activities of the brain (Schultz 1998). In order to clarify the relationship of these models to the current results, it is important to recall that the term *reinforcement learning* is used to describe a very large set of models (Dayan and Niv 2008). I-SAW is a member of this class of models. We believe that the most important difference between I-SAW and the popular reinforcement-learning model involves the error term discussed earlier. Like the SFP model, the popular reinforcement-learning models assume a normal error term. Other differences between I-SAW and the popular reinforcement-learning model involve the surprise trigger change assumption, and the abstraction of the recency effect. These new factors were introduced to capture the 6 phenomena summarized in Section 1.1 and are evaluated in the two choice-prediction competitions described in Section 1.3.

The naïve sampler model and probability matching. The naïve sampler model assumes random choice at the first trial and then reliance on a sample of size μ_i (property the agent) with replacement from all past experiences. It is an example of I-SAW with the constraints $\varepsilon_i = 0$, $w_i = 0$, $\rho_i = 0$, $\pi_i = 0$. The naïve sampler model captures the payoff variability effect, and underweighting of rare events, but cannot capture the other phenomena listed above. Best fit with these constraints is obtained with $\mu_i = 1, 2, 3, \ldots, 14$. The correlation between the aggregate choice rate and the model with these parameters is 0.83, and the MSD score is 0.019.

With the additional constraint $\mu_i = 1$, I-SAW implies the "probability matching" rule (matching the choice rate to the probability of success (see Estes 1950; Blavatskyy 2006). The correlation between the aggregate choice rate and the model with these parameters is 0.56, and the MSD score is 0.093.

1.1.9 Alternative Explanations of the Experience-Description Gap

Prospect theory (Kahneman and Tversky 1979; Wakker 2010), the leading model of decisions from description, captures three main behavioral regularities: overweighting of rare events, loss aversion, and the reflection effect (risk aversion in the gain domain and risk seeking in the loss domain). The results reviewed earlier show that different regularities emerge in the study of decisions from experience. The results reflect underweighting of rare events (Section 1.1.3), with no consistent indication for loss aversion (Section 1.1.2). In addition, under certain settings decisions from experience reveal a reverse reflection effect (Ludvig and Spetch 2011).

Recent research suggests several explanations for these differences. Our favorite explanation involves the assertion that decisions from description are a subclass of the larger class of decisions from experience. As in other subclasses, the decision makers tend to select strategies that have led to good outcomes in similar situations in the past. The experience-description gap emerges, under this explanation, as a result of two main effects of the available description. First, in certain cases, the description affects the set of strategies that can be used (see related ideas in Busemeyer and Myung 1992; Erev 1998; Rieskamp and Otto 2006; Erev and Roth 2001; Stahl 1996, 1999, 2000; Erev and Barron 2005). Second, the description affects the set of past experiences perceived to be similar.

In order to clarify the assertion that the description can affect the set of strategies, consider the following hypothetical choice problem:

Thought Experiment 1. Choose between

H	0 with certainty
L	\$1 with probability 0.99, −\$1,000,000 otherwise

It is easy to see that the availability of a description of the incentive structure will have a large effect here. Without a description (if this problem would be studied using the basic clicking paradigm), people are likely to select L at least until the first loss. With a description, reasonable individuals are expected to prefer H. This pattern can be captured with the hypothesis that the current description leads people to follow this rule: compute the expected values implied by the description, and select the best alternative based on this dimension. The apparent inconsistency between this hypothesis and the weak effect of description discussed in Sections 1.1.2 and 1.1.3 can be explained with the assertion that the tendency to use the EV rule decreases when the difference between the expected values, implied by the description, appear to be small and/or when the computation is too difficult (see Payne, Battman, and Johnson 1993). That is, the EV strategy is less likely to be used when the problem is similar to problems in which the EV rule was not found to be effective.

Marchiori, DiGuida, and Erev (2014) show that the current assertion can be used to explain "overweighting of rare events" in "one-shot decisions from description." Their explanation adds the assumption of overgeneralization from situations in which people decide based on subjective probability estimates. Subjective probability estimates tend to reflect overconfidence; for example, studies of probability estimates reveal that events estimated to occur 5% of the time actually occur about 20% of the time (Erev, Wallsten, and Budescu 1994). This overconfidence can be the product of random error: some of the events, estimated by probability 5%, occur with different probabilities, and this biases the occurrence rate toward 50%. Thus, the best reply to the belief that events that were estimated to occur with probability 5% will occur with higher probability tends to be reinforcing. Overgeneralization—from decisions with overconfident estimates to decisions under risk (when the described probabilities are accurate)—imply an initial overweighting of rare events in decisions under risk. Yet, experience eliminates this bias, and the tendency to rely on small samples can lead to the opposite bias.

Other likely contributors to the differences between the basic properties of decisions from experience and the predictions of prospect theory are presented shortly.

The white bear effect and the weighting of rare events. Wegner et al. (1987) note that when we try not to think about a white bear, a white bear comes to mind. This "white bear effect" can be one of the contributors to the tendency to overweight rare events in decisions from description. For example, it is possible that the gamble 5,000 with probability 1/1,000, and 0 otherwise seems attractive because we cannot avoid paying too much attention to the outcome 5,000 (see Birnbaum and Martin 2003). Underweighting of rare events in decisions from experience emerges, under this logic, because the availability of feedback reduces the attention given to the description and leads subjects to focus on the experienced outcome (Erev, Glozman, and Hertwig 2008).

Contingent loss aversion. The loss aversion assertion, one of the cornerstones of prospect theory (Kahneman and Tversky 1979), states that losses loom larger then gains. Thus, it predicts that when selecting among mixed prospects (prospects that can yield both gains and losses), people often prefer the safer prospect over riskier ones with higher expected value. The simplified investment problem examined in Section 1.1.2 reveals the opposite bias: a tendency to prefer the risky prospect even though the safe option provides higher expected return.

One explanation of this deviation from loss aversion is that it reflects a simple "experience-description gap" in the reaction to losses. This explanation is plausible, but it has two shortcomings. First, there are many situations in which people do not exhibit loss aversion in decisions from description (see Battalio, Kagel, and Jiranyakul (1990). Ert and Erev (2007, 2013), and the first trial in the simplified investment problem in Section 1.1.2). Most importantly, people appear to exhibit equal sensitivity to gains and losses in decisions from descriptions when the payoff magnitude is low. Thus, it is possible that small losses have a similar effect on decisions from experience and from description. And, the typical behavior in both cases reflects less loss aversion than implied by prospect theory (the predictions of prospect theory do not depend on the payoff magnitude).

A second shortcoming of the assumed experience-description gap in the reaction to losses is the observation that certain presentations of the outcomes can lead to behavior that appears to reflect loss aversion in decisions from experience (see Thaler et al. (1997) and a clarification in Erev, Ert, and Yechiam (2008)). For example, when people are asked to select between a "sure gain" or a risky prospect that provides higher expected return but often leads to a loss, they exhibit loss aversion when the payoffs are presented graphically (Thaler et al. 1997) but not when they are presented with clear numbers (Erev, Ert, and Yechiam 2008).

1.2 The Effect of Limited Feedback

Many natural decisions from experience problems involve situations in which the feedback is limited to the obtained payoffs. For example, when we choose to order a certain dish in a restaurant, we are not likely to know the outcome of ordering a different dish. The current section explores these decision problems with a focus on experiments that use the clicking paradigm (Figure 10.1) with limited feedback. That is, the feedback provided after each trial is limited to the outcome of the selected key.

Experimental studies that examine this set of limited-feedback situations highlight the generality of the six phenomena listed before. Yet, the results also demonstrate that the nature of the feedback can change the magnitude of the basic phenomena. The main changes can be described as reflections of the hot stove effect described next.

1.2.1 The Hot Stove Effect

Mark Twain (1897) asserts that after sitting on a hot stove lid, a cat is likely to avoid sitting on stove lids even when they are cold. Denrell and March (2001; also see Denrell (2005, 2007) and a related observation in Einhorn and Hogarth (1978)) show that Twain's assertion is a likely consequence of learning when the feedback is limited to the obtained payoff. Learning in this setting increases risk aversion. This observation, referred to as the hot stove effect, is a logical consequence of the inherent asymmetry between the effect of good and bad experiences. Good outcomes increase the probability that a choice will be repeated and for that reason facilitate the collection of additional information concerning the value of the alternative that has yielded the good outcome. Bad outcomes reduce the probability that the choice will be repeated and for that reason impair the collection of additional information concerning the value of the alternative that has yielded the bad outcome. As a result, the effect of bad outcomes is stronger (lasts longer) than the effect of good outcomes. Since options with a high variability are more likely to produce bad outcomes, the hot stove hypothesis predicts a decreasing tendency to choose such options.

One indication of the descriptive value of the hot stove effect is provided by a comparison of choice behavior with and without foregone payoffs in the four-alternative Iowa gambling task discussed earlier. The availability of foregone payoffs tends to increase risk taking (see Yechiam and Busemeyer 2006). A similar pattern was documented by Fujikawa (2009) in an analysis of problem 9. His analysis suggests that the hot stove effect can reduce underweighting of unattractive rare events.

Additional experimental studies demonstrate that the magnitude of the hot stove effect is maximal when the risky alternative is a long-shot gamble. Table 10.2 illustrates this pattern. It presents the proportion of R choices in 12 problems that were run for 100 trials using the basic clicking paradigm (with complete feedback), with and without forgone payoffs (the limited feedback conditions were run by Erev et al. (2010a), and the complete feedback conditions were run by Nevo and Erev (2012)). The results (presented in two blocks of 50 trials) reveal a large hot stove effect in "rare treasure" problems, when the probability of a high payoff from risky choice is 0.1 or lower: In all seven problems of this type, choice of the risky alternative in the last block is higher in

TABLE 10.2:
The Proportion of risky choices as a function of feedback and time (in two blocks of 50 trials) in the games studied in Nevo and Erev (2012) and Erev et al. (2012a).

Problem		*Block*	*Complete*	*Partial*	*Difference*
14	S 7 with certainty	1	0.45	0.21	0.24
	R (16.5, 0.01; 6.9)	2	0.46	0.15	0.31
15	S −9.4 with certainty	1	0.27	0.16	0.11
	R (−2, 0.05; −10.4)	2	0.25	0.07	0.18
16	S −4.1 with certainty	1	0.51	0.31	0.20
	R (1.3, 0.05; −4.3)	2	0.58	0.29	0.29
17	S −18.7 with certainty	1	0.37	0.35	0.02
	R (−7.1, 0.07; −19.6)	2	0.39	0.33	0.06
18	S −7.9 with certainty	1	0.41	0.24	0.17
	R (5, 0.08; −9.1)	2	0.49	0.14	0.35
19	S −25.4 with certainty	1	0.29	0.11	0.18
	R (−8.9, 0.08; −26.3)	2	0.32	0.07	0.25
20	S 11.5 with certainty	1	0.31	0.12	0.19
	R (25.7, 0.1; 8.1)	2	0.28	0.11	0.17
21	S −15.5 with certainty	1	0.65	0.62	0.03
	R (−8.8, 0.6; −19.5)	2	0.71	0.69	0.02
22	S 2.2 with certainty	1	0.48	0.52	−0.04
	R (3, 0.93; −7.2)	2	0.46	0.35	0.11
23	S 25.2 with certainty	1	0.54	0.65	−0.11
	R (26.5,0.94; 8.3)	2	0.49	0.42	0.07
24	S 6.8 with certainty	1	0.54	0.70	−0.16
	R (7.3, 0.96; −8.5)	2	0.47	0.60	−0.13
25	S 11 with certainty	1	0.61	0.69	−0.08
	R (11.4, 0.97; 1.9)	2	0.53	0.63	−0.10

the complete feedback condition. The pattern in the 5 problems with higher probability for the high payoff from risky choice is less clear.

Diminishing exploration. As noted before, the hot stove effect is implied by all models that assume positive recency. Explaining the interaction of the observed effect with time and with the magnitude of the high payoff from the risky option is more challenging. The most natural explanation for the effect increasing with time (from the first to the second block) can be captured with the assertion of diminishing exploration: a high exploration rate in the beginning of the experimental session and a lower rate of exploration toward the end.

The observation that the hot stove effect was not observed in problems where the risky prospect leads to better outcomes most of the time can be the product of the fact that even limited exploration is enough, in these cases, to demonstrate the value of the risky option. If some exploration continues even after an extreme low payoff, the hot stove effect is not likely to emerge in these settings.

Two-armed bandit problems. The task faced by the subjects in the limited feedback conditions of summarized in Table 10.2 is similar to the 2-armed bandit problem (see

Degroot 1970; Gittins 1989). Yet, the common analyses of 2-armed-bandit problems focus on situations in which the decision makers have more information and the optimal strategy can be computed. Specifically, decision makers know the expected payoff from the safe option and know that the risky option provides one of two outcomes with fixed probability. Theoretical analysis of these 2-armed-bandit problems shows that the optimal strategy is to start with exploration of the risky option and switch to the safe option if the outcome fall below a certain cutoff. Thus, the diminishing exploration pattern suggested here is similar to the optimal strategy in these simpler problems.

Direct experimental studies of 2-armed-bandit problems show the robustness of the pattern discussed earlier. Meyer and Shi (1995) show an increase in counterproductive exploration with payoff variability and a slow reduction in exploration (not enough exploration in the beginning and too much exploration in the end). Gans, Knox, and Croson's (2007) results suggest large individual differences and a "very recent" effect.

1.2.2 I-SAW with Diminishing Exploration

Erev, Ert, and Yechiam (2008) show that the main properties of binary decisions from experience with limited feedback can be captured with an "exploration sampler" model that assumes reliance on small samples and diminishing exploration. The main assumptions of this model can be captured in an extension of I-SAW (Section 1.1.7) that adds the assumption that the probability of exploration depends on the available feedback. I-SAW assumes that when the feedback is complete (includes information concerning obtained and forgone payoff), the probability of exploration is fixed during the experiment and reflect an individual trait (ε_i). The extended version adds the assumption that when the feedback is limited to the obtained payoff, the probability of exploration starts at 1 and diminishes with time. The speed of the decline in exploration is assumed to depend on the expected length of the experiment. Specifically, we assume that in this case the probability of exploration at trial t equals $\varepsilon_i^{(t-1)/T}$, where T is the length of the experiment (in the experiments reviewed in Table 10.2, $T = 100$). In addition, the extension of I-SAW to situations with limited feedback implies that less information is used during sampling and during the computation of surprise outcomes: only the obtained payoffs are used.

1.3 Two Choice-Prediction Competitions

We believe that the basic learning phenomena just considered are an important part of the factors that shape human behavior. This optimistic belief implies that good models of the joint effect of these phenomena can provide useful ex ante predictions of the effect of economic incentives in a wide set of situations (Erev and Roth 1998). Two choice-prediction competitions that evaluate this optimistic hypothesis and facilitate the comparison of alternative learning models are described below.

1.3.1 The Technion Prediction Tournament: Individual Decisions with Limited Feedback

Erev et al. (2010a) present a choice-prediction competition designed to facilitate the development and comparison of models of decisions from experience under limited feedback.[8] The organizers of the competition (the first three coauthors of that paper) ran two large experimental studies using the clicking paradigm without information concerning forgone payoffs. Each study focused on 60 randomly selected problems. All the problems involved a choice between a safe prospect that provides a medium payoff (referred to as M) with certainty and a risky prospect that yields a high payoff (H) with probability Ph and a low payoff (L) otherwise. Thus, the basic choice problem is:

S: M with certainty
R: H with probability Ph; L otherwise (with probability 1 − Ph)

The four parameters (M, H, Ph, and L) were randomly selected with a well-defined algorithm that implies (1) the possible payoffs were between −30 and +30 shekels (1 shekel equaled about $0.30); (2) L < H; (3) M was between L and H in 95% of the problems; and (4) the difference between the expected values of the two prospects was relatively small. Twelve of the 120 problems that were examined are presented in Table 10.2.

The first study, referred to as the estimation experiment, was run in March 2008. Each of the 60 problems was faced by 20 subjects for 100 trials. Each subject played 12 games, and the payoffs (in shekels) were determined by a randomly selected trial. In April 2008, the organizers posted the results and the best baseline models that they could find on the Web (see http://tx.technion.ac.il/∼erev/Comp/Comp.html) and challenged other researchers to predict the results of the second study. The second study, referred to as the competition experiment, was run in May 2008 using the same experimental method as the estimation experiment but different randomly selected problems and different subjects. The results of the competition study were not revealed until September 2, 2008.

Researchers participating in the competitions were allowed to study the results of the estimation study. Their goal was to develop a model that would predict the results (the mean choice proportion over all choices in each problem) of the competition study. The model had to be implemented in a computer program that reads the payoff distributions of the relevant gambles as an input and predicts the proportion of risky choices as an output. The submission deadline was September 1, 2008. The submitted models were ranked based on the mean squared deviation (MSD) between the predicted and the observed choice proportions.

ENO (equivalent number of observations). One advantage of the MSD criteria used here is its relationship to traditional statistics (like regression, t-test and the d-statistic) and its intuitive interpretation. These attractive features are clarified with the computation of the ENO (equivalent number of observations) order-maintaining transformation of the MSD scores (see Erev et al. 2007). The ENO of a model is an estimation of the size of the experiment that has to be run to obtain predictions that are more accurate than the model's prediction. For example, if a model's prediction of the probability of risky choices in a particular problem has an ENO of 10, this prediction is expected to be as accurate as the prediction based on the observed proportion of risky choices in an experimental study of that problem with 10 participants.

Results. The models evaluated in the competition can be classified in two main classes: the first includes instance-based models like I-SAW that assume that the agents remember specific experiences (and tend to rely on small samples). The second includes models that do not assume memory of and/or reliance on specific experiences. About half of the baseline models and half of the submissions belong to each class. The results reveal a large advantage of the instance-based models. The best baseline model was a predecessor of I-SAW. The ENO of this best baseline was 47.2. In the current context, the predictions of this model are almost identical to the predictions of the refined model, I-SAW, with the parameters $\varepsilon_i \sim U[0, 0.20]$, $w_i \sim U[0, 1]$, $\rho_i \sim U[0, 0.6]$, $\pi_i \sim U[0, 0.6]$, and μ_i drawn from integers 1 to 14.[9]

The winner of the competition was an instance-based model that assumes an ACT-R cognitive architecture.[10] Like the best baseline and I-SAW, the winning model builds on the instance-based learning model proposed by Gonzalez, Lerch, and Lebiere (2003) and implies reliance on small samples from experience. The winner had slightly lower ENO (32.5) than the best baseline model (the baseline models did not participate in the competition), with two attractive features. First, the ACT-R cognitive architecture involves a psychologically more realistic abstraction of the relevant memory processes. For example, it assumes a continuous weighting of all past experiences. Second, the winning ACT-R model is rather general; it was designed to capture decisions in static as well as dynamic environments. We return to this point later.

Analysis of the predictions of the models in the competition that do not assume memory of specific experience suggests that their most important failure involves the effect of Ph (the probability of high payoff from risky choice). With the parameters that best fit the data, these models underpredict the R-rate (risk taking). That is, these models overpredict the hot stove effect. This pattern results from extremely low payoffs from the risky prospect decreasing the probability of exploring that prospect. Recent research (Shteingart, Neiman, and Loewenstein 2013) shows that this shortcoming of reinforcement-learning models that do not store specific instances can be addressed by assuming oversensitivity to the very first experience. Their model implies reliance on a very small sample, without explicit memory of this experience.

Another outcome from the competition involves the estimation technique: all the leading submissions used a computer-simulation-based estimation method and did not use more sophisticated, one-period-ahead, econometric techniques. This is surprising, as previous research shows that when the model is "well-specified," the correct one-period estimation provides the best estimate of the parameters. One explanation for this is that current models are misspecified, and the one-period-ahead techniques are more sensitive to this misspecification (see Erev and Haruvy 2005).

1.3.2 The Market Entry Game Competition: Social Interaction with Complete Feedback

Erev et al. (2010b) organized a choice prediction competition that focuses on 4-person market entry games under limited prior information. The experimental subjects were informed that they would play a market-entry game and have to select between a risky entry to the market and a safer decision to stay out of the market.

The payoffs depended on a realization of a binary gamble (the realization at trial t is denoted G_t, and yields H with probability Ph; and L otherwise), the number of entrants (E), and two additional parameters (k and S). The exact payoff for player i at trial t was

$$V_i(t) = \begin{cases} 10 - k(E) + G_t & \text{if } i \text{ enters} \\ \text{round}(G_t/S) \text{ with } p = 0.5; -\text{round}(G_t/S) \text{ otherwise} & \text{if } i \text{ does not enter} \end{cases}$$

The parameters H, Ph, L, k and S were randomly drawn under certain constraints (e.g., the expected value of the gamble was zero and the mean entry rate at equilibrium was 0.5).

The participants did not receive a description of the payoff rule and had to rely on complete feedback (obtained and forgone payoffs) after each trial. The organizers ran an estimation study with 40 games and a competition study with 40 additional games.

The results of the estimation study were published in May 2010, and the submission deadline was September 2010. Analysis of the estimation study showed that the results exhibit the basic learning phenomena documented in the individual choice tasks summarized in Section 1.1. In addition, the result shows a high initial entry rate: 66% in the first trial. Comparison of several baseline models highlights the advantage of I-SAW over other models. Best fit was provided with a slight modification of the “strategy set simplification assumption”: the best baseline model is I-SAW with the added assumption of an initial tendency to enter the market in 66% of the trials.

Twenty-five teams participated in the competition. The submitted models included basic reinforcement learning, neural networks, ACT-R, and I-SAW-like sampling models. The results reveal a large advantage of instance-based models that assume reliance on small samples and surprise-triggers-change. Indeed, all 10 leading submissions belong to this class of models. The winner of the competition (Chen et al. 2011) is a variant of I-SAW that adds the assumption of bounded memory. The runner-up (Gonzalez, Dutt, and Lejarraga 2011) is a refinement of the instance-based learning model proposed by Gonzalez, Lerch, and Lebiere (2003).

The ENO of I-SAW (in predicting the average payoff, a statistic that captures the entry rate and implied coordination level) in the last block of 25 trials was 42.2. As in the first competition, traditional “normal error term” reinforcement-learning models that do not assume reliance on specific instances did not do well. It seems that the main reason for their failure involves the coexistence of underweighting of rare events and a relatively weak recency effect. The traditional reinforcement learning models (and similar fictitious play and experience weighted attraction models; Camerer and Ho (1999)) that were evaluated have to assume a strong recency effect in order to capture the observed underweighting of rare events.

Another similarity to the first competition involves the estimation techniques used by the best models. All the top submissions used simulation-based methods and avoided more sophisticated one-period-econometrics.

2 DYNAMIC ENVIRONMENTS

Many of the early experimental studies of learning focused on the effect of training in one environment (the training phase) on performance in another environment (test phase). Thus, they examined decisions in dynamic environments. Some of the classical results documented in these settings are reviewed next.

2.1 *The Partial Reinforcement Extinction Effect and Reinforcement Schedules*

The partial reinforcement extinction effect (PREE) is one of the best-known phenomena documented in classical behavioral research. The effect implies that under partial reinforcement schedules (where some responses are, randomly, not reinforced), learned behavior is more robust to extinction, in comparison to continuous reinforcement. This effect was first demonstrated in Humphreys’ (1939a) examination of eye blinks in rabbits.

Humphreys (1939b) and Grant, Hake, and Hornseth (1951) show PREE in human behavior. These studies focused on “predicting whether a lightbulb will flash or not.” Participants were presented with two lightbulbs. On each trial, the right-hand bulb was blinking, and the participants had to predict whether the left bulb would blink as well.

The classical experiments included training and extinction phases and compared two conditions: continuous reinforcement and partial reinforcement. The two conditions differ during the training phase: The response yes (i.e., the prediction that the left light bulb would flash) was reinforced on 100% of the trials under continuous reinforcement and in only some of the trials under partial reinforcement. In the extinction phase, yes was never reinforced. The results demonstrated that in the extinction phase, the rate of yes responses decreased faster for the continuous-reinforcement schedule group than for the partial-reinforcement schedule group. However, during training, learning was faster as the reinforcement rate increased.

Hochman and Erev (2013) replicated the PREE using the clicking paradigm. One of their studies focused on the following problems:

Problem 26—continuous (r = 100, n = 11, FB = complete, 1 point = ¢ 0.25)

S	8 with certainty
R	9 with certainty

Problem 27—partial (same procedure as in Problem 26)

S	8 with certainty
R	17 with probability 0.5; 1 otherwise

Problem 28—extinction (same procedure as in Problem 26)

S	8 with certainty
R	1 with certainty

The study included two phases, acquisition (the first 100 trials) and extinction (the last 100 trials). During the acquisition phase, one group of participants (the continuous group) played problem 26, and the second group (the partial group) played problem 27. During the extinction stage, option R was dominated: Both groups were faced with problem 28 at this phase. The participants were not informed that the experiment included two phases.

The results (cf. Figure 10.5) reveal more R choices in the continuous group during the acquisition phase and the opposite pattern during the extinction phase. Thus, payoff variability slows the initial learning for R choices, but it also slows the extinction of this behavior.

Hochman and Erev (2013) show that the PREE pattern they observed can be captured with a variant of I-SAW that adds the assumption that perceived similarity is determined by the sequence of the last 4 outcomes from R. In order to clarify the intuition behind this observation, consider the decision at trial 106 after the payoff sequence 1, 1, 1, 1 from R. The participants in the continuous group saw this pattern only once in the past (at trial 105), and the outcome from R in that case was disappointing (R gave 1, and S paid 8). Thus, they are predicted to select S. In contrast, the participants in the partial group have seen this sequence several times during the first 100 trials, and in some of these cases it was followed by high payoff from R (17); thus, depending on their exact sample, they may choose R.

Problem	S	R
26	8 with certainty	9 with certainty
27	8 with certainty	(1, 0.5; 17)
28	8 with certainty	1 with certainty

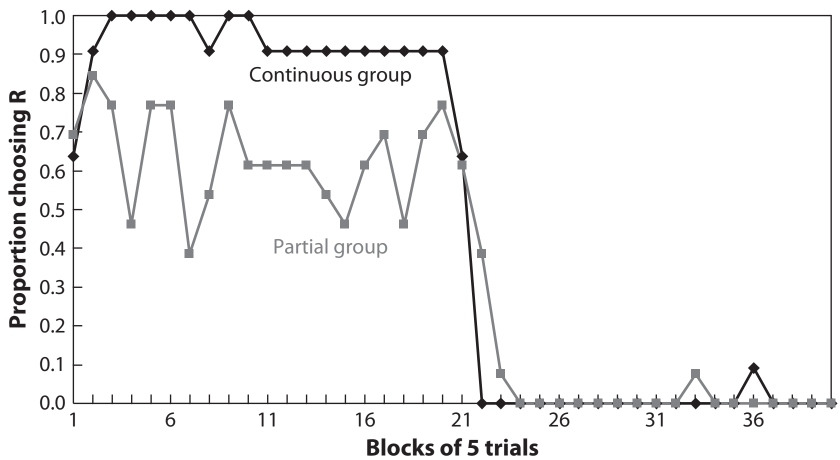

Figure 10.5: The partial reinforcement extinction effect. Continuous group played problem 26 and then 28; partial played 27 and then 28.

2.2 Spontaneous Alternation, the Gambler Fallacy, and Response to Patterns

Tolman (1925) observed an interesting violation of the law of effect in a study of rats' behavior in a T-maze. Upon receiving a reinforcement in a particular arm, rats tend to switch to the other arm of the maze. According to the common explanation of this spontaneous alternation pattern (see the review in Dember and Fowler 1958), it reflects a tendency to respond to the likely sequential dependencies in natural settings. That is, in most environments where rats eat (e.g., storehouses and garbage dumps), food is replenished independently of feeding. Thus, after eating the food in one location, it is typically optimal to move to a different location.

More recent studies use a similar argument to explain probability matching (see Estes 1976; Sonsino 1997; Gaissmaier and Schooler 2008) and underweighting of rare events (Plonsky, Teodorescu, and Erev 2015). These phenomena can be a result of an effort to respond to patterns and sequential dependencies in the environment that implies reliance on small samples. When the environment is static and noisy, this effort impairs maximization. When the environment changes in a consistent fashion, however, sensitivity to sequential dependencies can be very useful (e.g., Gonzalez et al. 2003; Sterman 1989). One example of effective adaptation to consistent change is provided by the continuous condition in the PREE studies (e.g., the change from problem 26 to problem 28). Gaissmaier and Schooler show that people can respond to consistent patterns even when the detection of the pattern requires sensitivity to the last 12 outcomes.

2.3 *Negative and Positive Transfer*

The effect of learning in one task on the performance of a different task is referred to as *transfer*. Transfer is highly sensitive to the characteristics of the two tasks (see Osgood (1949) and analysis of economic implications in Cooper and Kagel (2003)). Whereas many studies document positive transfer (improved performance on the second task), other studies document no transfer and even negative transfer. Moreover, many studies report both negative and positive transfer in the same setting. One example is provided by the transfer from problem 26 to 28: the initial transfer in this case is negative (less than 50% maximization rate in the first few transfer trials), but the long-term effect is positive (higher maximization rate in problem 28 when it is played after problem 26).

One explanation for the existence of positive and negative transfer involves the assertion that people learn cognitive strategies (rather than situation-specific actions). For example, in problem 26 they might learn to prefer "best reply to recent experiences" over "alternation." This learning leads to negative transfer in the first trials of problem 28 (S-rate below 50%) but to positive transfer after sufficient experience with problem 28 when recent experience implies that S leads to better outcomes.

2.4 *The Effect of Delay and Melioration*

Thorndike (1911) demonstrates that behavior is highly sensitive to the timing of the reinforcement. Delay of the reinforcement slows learning. This tendency implies (see Kagel, Battalio, and Green 1995) that animals behave as if they prefer a smaller immediate reward to a larger delayed reward and that this preference is not consistent with a simple discounting explanation. A clear demonstration of this pattern is provided by Green et al. (1981) in a study that used a variant of the clicking paradigm.

Each trial consisted of a 30-second choice period, during which a pigeon was presented with a choice between two keys, followed by an outcome. One key led to a small reward—2 seconds of access to a grain hopper with a delay of x seconds—and the other to a larger reward—6 seconds of access to a grain hopper, with a delay of $x+4$ seconds. The time variable x varied from 2 to 28 seconds.

The results reveal that when x is low (less than 5 seconds), each bird strongly favored the smaller, more immediate outcome. The nearly exclusive preference for the smaller reward means that the pigeons failed to maximize total food intake. However, as the delay between choice and both outcomes (the time x) increased, preference reversed, with nearly every bird now choosing the larger, more delayed outcome on more than 80% of the trials. That is, with longer delays the pigeons maximized total food intake.

Melioration. Herrnstein and his associates (Herrnstein 1988; Herrnstein and Vaughan, 1980; Herrnstein and Mazur 1987; Herrnstein and Prelec 1991) demonstrate that in certain settings the tendency to underweight delayed payoff can lead to a robust deviation from maximization. Specifically, they show that experience can lead decision makers to behave as if they meliorate (maximize immediate payoffs) rather than to maximize long-term expected utilities.[11]

For a simple demonstration of this regularity using the clicking paradigm, consider the following choice task.

Problem 29 ($r=200$, $n=20$, FB = complete, 1 point = 0.01 shekel)

S	1 with certainty	[S-rate: 90%]
R	+10 points with $p = N(\mathrm{R})/(50+t)$; 0 otherwise	

Here t is the trial number and N(R) is the number of R choices made by the participant before trial t. It is easy to see that if the experiment is long enough, option R maximizes long term expected payoff. Yet, melioration implies S choices.

The data for problem 29 in a 200-trial experiment reveal strong support for the melioration hypothesis. The choice rate of option S (melioration) over the last 100 trials was 90. All 20 subjects chose S on more than 50% of the trials.

Herrnstein et al. (1993) show that melioration decreases with clear information concerning the long-term effect of available choices. Thus, the evidence for melioration is best described as indicative of insufficient exploration.

2.5 *Models of Learning in Dynamic Settings*

Gonzalez, Lerch, and Lebiere (2003) show that the main properties of decisions from experience in dynamic settings can be captured with a variant of the ACT-R model (see Anderson and Lebiere 1998) that assumes similarity-based weighting of all relevant experience. Under this model, decision makers are assumed to overweight a small set of experience that occurred in situations that seem most similar to the current setting and give lower weight to other experience. As noted before, this idea was also found to capture behavior in static settings: It is the basis of the instance-based model that won the choice- prediction competition described in Section 1.3.1.

A shortcoming of the similarity-based approach is the determination of a similarity function. Different studies appear to support different similarity functions. For example, Gonzalez et al. show an important role for temporal similarity (also see Hochman and Erev (2013) and Section 2.1) and that this is best determined by the sequence of the last four outcomes. Plonsky, Teodorescu, and Erev (2015) suggest that these apparent inconsistencies can be a reflection of the fact that people consider a wide set of similarity functions, and try to select the best function. When the environment is highly dynamic and predictable, the probability of success is high. However, when the environment is noisy, the probability of success is low, and the observed behavior can be approximated by relying on small samples of randomly selected past experience (recall Section 1.1.7).

Recent research shows that learning in dynamic setting can also be captured with reinforcement-learning models that include a recognition process that categorizes cues into situations (see Redish et al. 2007). Gershman , Blei, and Niv (2010) refine this observation and show the value of Bayesian inference within a reinforcement-learning model that assumes an unbounded number of latent causes.

3 MULTIPLE ALTERNATIVES AND ADDITIONAL STIMULI

Unlike the simple binary-choice clicking experiments reviewed earlier, most natural activities involve learning among multiple alternatives based on multiple sources of information. Even in the road-crossing example, the decision maker can choose between many actions (different ways to cross the road and different alternatives to this behavior) and can use many signals. Experimental studies that explore learning among multiple alternatives and the effect of different signals are reviewed next.

3.1 *Successive Approximations, Hill Climbing, and the Neighborhood Effect*

Skinner (1938) highlights the value of the method of successive approximations (also known as “shaping”) for teaching complex behavior. Shaping is used when the desired behavior is not observed initially. The procedure involves first reinforcing some observed behavior only vaguely similar to the one desired. Once that behavior is established, the trainer looks for (reinforces) variations that come a little closer to the desired behavior, and so on. Skinner and his students have been quite successful in teaching simple animals to do some quite extraordinary things. For example, they taught a pigeon to control a missile (Glines 2005).

The basic idea behind the method of successive approximations is the assumption that there are many strategies that can be used in an attempt to perform a complex task. That is, the set of feasible strategies is very large. The agent tends to consider strategies similar to the reinforced strategies. As a result, learning does not ensure convergence to the globally optimal strategy. It can lead to a local optimum. The method of successive approximations is effective because it reduces this risk (at least when the trainer has a good understanding of the location of the optimal strategy).

A clear demonstration of the tendency to converge to a local optimum is provided by Busemeyer and Myung's (1988) examination of choice behavior in a multiple-alternative resource-allocation task. In each trial the participants were asked to divide limited resources among three issues. Each allocation can be abstracted as a selection of one of many possible allocations (strategies) that can be placed in a triangle (called the simplex). The results reveal that performance is highly sensitive to the location of the different strategies in the simplex. Higher maximization rate was observed when the best strategies were in the same "neighborhood." Busemeyer and Myung note that this regularity can be captured by a hill-climbing search process.

Erev and Barron (2002) replicated this observation in a study that focused on problems 30 and 31 using the clicking paradigm with limited feedback. Both problems involve a choice among the same 400 alternatives. Each alternative is associated with only one outcome. The two problems differ with respect to the location of the 400 alternatives in the 20 ×20 matrix presentation. The top panel in Figure 10.6 shows a three-dimensional summary of the two matrices. It shows that both matrices have two maximum points (a local maximum of 32 and a global maximum of 52). The conversion rate was 0.25¢ per point. In Problem 30 the local maximum (32) had a wide basin of attraction. Problem 31 was created by swapping the location of the two maxima; thus, the global maximum (52) had the wide basin of attraction.

The lower panel in Figure 10.6 presents the proportion of maximization under the two conditions. In line with Busemeyer and Myung's findings, the decision makers were closer to maximization in problem 31 (global maximum with wide basin of attraction) than in problem 30. Since maximization rate seems to depend on the relative location of the global maximum, we refer to this result as the neighborhood effect. Yechiam, Erev, and Gopher (2001) clarify the relationship between convergence to a local optimum and shaping. They show that a minimalistic shaping procedure, the prevention of repeated choice, reduces the tendency to converge to a local maximum in a variant of problem 30.

Implications to descriptive models. The attempt to model learning among multiple alternatives given incomplete feedback highlights the importance of the details of the assumed exploration process. Busemeyer and Myung (1988) show that the main features of the exploration process can be captured with a hill-climbing rule. Erev and Barron (2002) and Yechiam, Erev, and Gopher (2001) show the value of modeling hill climbing as one of several cognitive strategies. The model assumes reinforcement learning among these strategies. Rieskamp et al. (2003) highlight the value of a model that assumes a focus on the difference between the current results and the best past experience.

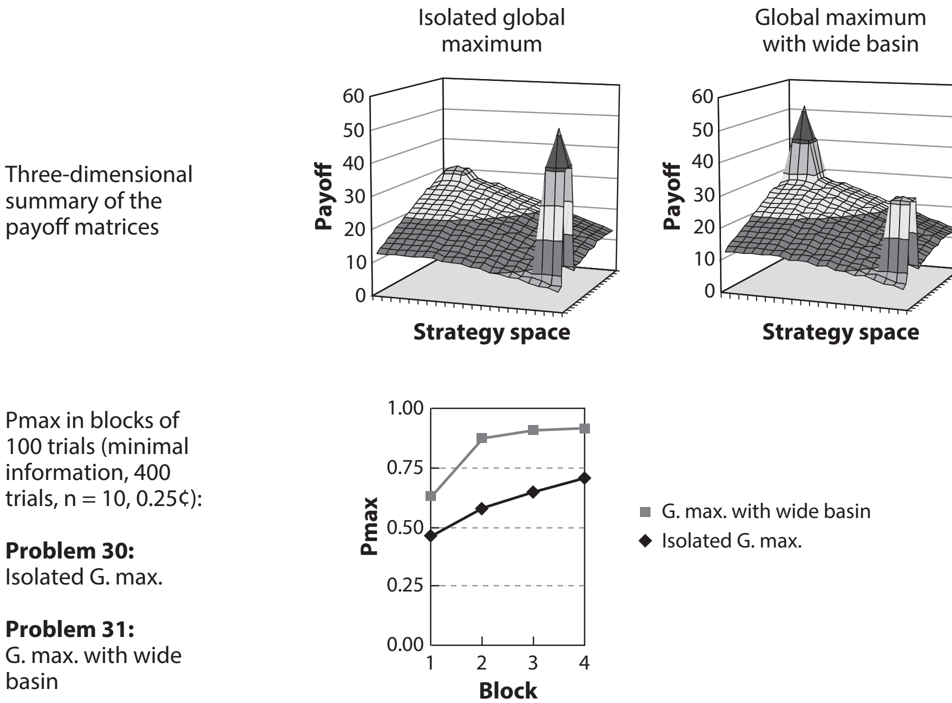

Figure 10.6: The top panel represents the payoff matrices. The lower panel presents the proportion of the maximal payoff (Pmax) in 4 blocks of 50 trials.

Analysis of exploration by firms (Levinthal and March 1993; Gavetti and Levinthal 2000) highlights the value of a distinction between two types of exploration: forward looking and backward looking. Teodorescu and Erev (2014a) demonstrate that this distinction can also shed light on individual choice behavior among multiple alternatives using the clicking paradigm. Their results reflect insufficient exploration in "rare treasure problems" (when the common outcome of exploration is disappointing), and overexploration in a rare mines problem (when the common outcome of exploration is attractive). The coexistence of under- and overexploration can be captured with an extension of I-SAW that assumes a choice between cognitive strategies (exploration or exploration) before the choice between the actions.

3.2 Learned Helplessness

Overmier and Seligman (1967) found that dogs exposed to inescapable shocks in one situation later failed to learn to escape shock in a different situation where escape was possible. Follow-up research (see the review in Maier and Seligman 1976) shows that this "learned helplessness" phenomenon is robust across species and experimental paradigms and provides an insightful account of human depression.

Teodorescu and Erev (2014b) replicated the learned helplessness pattern in the clicking paradigm and compared three explanations for the results. The three explanations differ with respect to the assumed cause for the tendency to give up too early (and exhibit insufficient exploration). The trigger can be (1) the belief that environment is uncontrollable, (2) low average reinforcement from exploration, and (3) low probability of success. The results favor the third explanation.

3.3 *Multiple Alternatives with Complete Feedback*

An increase in the number of possible alternatives increases the importance of the availability of information concerning the forgone payoffs. When the payoff variability is low, the availability of complete feedback facilitates maximization and leads to very quick learning to prefer to the best option (Grosskopf et al. 2006). However, when the payoff variability is large, the availability of complete feedback can lead to the big eyes effect (see Section 1.1.2) that can impair maximization.

Ert and Erev (2007) examined a 50-alternative problem (using the clicking paradigm with complete feedback that included the payoff from all 50 alternatives after each choice) in which the predictions of the big eyes effect contradict the predictions of underweighting of rare events. Half of the 50 alternatives provided 3 with certainty, and the other half provided 32 in 10% of the trials and 0 otherwise. Thus, the risky option maximized expected value, and the big eye effect implies risky choice (because the best outcome over the 50 alternatives tends to be 32 from one of the risky alternatives). The choice rate of the risky option (after 50 trials) was only 40%. It seems that in the current setting, underweighting of rare events is stronger than the big eyes effect. This pattern can be captured with the assertion that regret reduces payoff sensitivity. Another explanation assumes limited attention. Specifically, it is reasonable to assume that when the number of alternatives is very large, people cannot attend to all the forgone payoffs (see a related idea in Camerer and Ho 1999).

3.4 *The Effect of Additional Stimuli (Beyond Clicking)*

The current review focuses on the direct effects of obtained and forgone payoffs on choice behavior. We believe that these effects are the most important drivers of human adjustment to economic incentives. Yet, in certain settings other factors can affect this adjustment process. Two important examples are discussed next.

3.4.1 Pavlovian (Classical) Conditioning

The early psychological study of learning distinguishes between two classes of basic processes: instrumental and Pavlovian conditioning. Instrumental conditioning (also known as operant conditioning) describes behavior in situations in which the agent learns to prefer specific voluntary actions that affect the environment. Thus, all the studies summarized earlier are examples of instrumental conditioning.

The early definition of Pavlovian conditioning focuses on the association between two stimuli. For example, in each trial of Pavlov's (1927) classical study, dogs were presented with a bell a few seconds before receiving food. At the beginning of the study, the bell elicited no response, and the food elicited salivation (unconditioned response, UR). After several trials the dogs started salivating immediately after hearing the bell. Thus, the bell is called a conditioned stimulus (CS), and the food is called an unconditioned stimulus (US).

At first glance Pavlovian conditioning does not appear to be very important in the analysis of economic behavior. However, Rescorla and Solomon (1967) show that more careful analysis can lead to different conclusions: Since Pavlovian conditioning determines emotion and related innate states, it is natural to assume that it affects the subjective interpretation of the choice environment. Rescorla and Solomon (see related ideas in Mowrer 1947) propose a two-process model that captures this idea. Under this model, instrumental conditioning drives learning in each subjective state, but Pavlovian conditioning determines the subjective state. Since agents are likely to learn different behavior in different subjective states, Pavlovian conditioning can be highly important.

One example of the importance of the subjective states is provided by the dynamic task considered in Section 2, in which the payoff rule changed between the first 100 trials and the last 100 trials (i.e., the payoff rule changed from problem 26 to problem 28) without the subjects being instructed of this two-phase structure. In this setting, distinguishing between the different objective states of the world enhances performance. Thus, if Pavlovian conditioning determines the agent's responsiveness to these and similar states, it determines in part the learning process.

It is interesting to note that Rescorla and Solomon's theory implies a very different effect of emotions than the common abstraction in economic models of emotion. Under the common abstraction (e.g., Fehr and Schmidt 1999; Bolton and Ockenfels 2000), emotions like inequality aversion affect subjective utility. For example, people reject unfair offers in the ultimatum game because the rejection reduces disutility (negative emotion) from inequality (Fehr and Schmidt 1999; Bolton and Ockenfels 2000). Rascorla and Solomon's analysis can be used to support the assumption that the main effect of emotion involves the generalization from specific past experiences. In other words, rejection of unfair offers may be a product of an emotion that directs the agent to select a behavior learned in an environment in which rejection of unfair offers is adaptive (see related observations in Cooper and Kagel (Chapter 4).

Another example of the economic implications of Pavlovian conditioning involves addiction. Smith and Tasnádi (2007) show that "harmful" addiction can be the result of a mismatch between behavioral (learning) algorithms encoded in the human genome and the expanded menu of choices faced by consumers in the modern world.

3.4.2 Observational Learning

Observational learning refers to learning by observing others' decisions and payoffs. A number of animal studies support observational learning. Terkel (1996) shows that young rats learn to skin pinecones by observing their mothers. John et al. (1969) show that cats can learn tasks by observing the performance of an animal already trained in that particular task.

Miller and Dollard (1941) argued that observational learning is no different than simple reinforcement learning in that observational learning involves situations where the stimulus is the behavior of another person and the payoff maximizing behavior happens to be a similar behavior. In one of their experiments, first-grade children were paired, with one in the role of leader and the other in the role of follower. In each trial, the children sequentially entered a room with two boxes. In one of the boxes, there was candy. The leader first chose a box and obtained any candy that was in it. The follower observed which box the leader chose but not the outcome of that choice. Next, the contents of the boxes were emptied and candy was again placed in one box. The placement of the candy was manipulated in two treatments. In one treatment, the candy was placed in the box previously selected by the leader. In the other treatment, candy was placed in the box not chosen by the leader. The follower then entered the room and chose a box. After a few trials, children in the first group always copied the response of the leader and children in the second group made the opposite response.

Bandura (1965) argued that the payoff received by the observed person should matter in the decision of whether to imitate that person. In Bandura, a group of 4-year-old children watched a short film on a TV screen in which an adult exhibited aggressive behavior toward an inflated "bobo doll." The children then saw the aggressor being reinforced by another adult. In one treatment, the aggressor was praised and given soda and snacks. In a different treatment, the adult was scolded, spanked, and warned not to do it again. The children were then left in a room with the doll, along with other toys. The imitation and aggression were more pronounced when the adult was observed receiving a reward for his actions and less pronounced when the adult was punished.

Merlo and Schotter (2003) raise the prospect that in some settings observational learners may learn better than subjects engaged in the task. In their experiments, subjects chose a number between 0 and 100. The higher the number chosen, the higher the cost incurred by the subject and the higher the probability of winning the high prize, resulting in an interior optimal choice of 37. Subjects in the baseline experiment repeated the decision task 75 times and were paid a small amount after each trial. As each subject performed the experiment, another subject watched over his or her shoulder. In the end of the 75 trials, the observers as well as the active subjects were both given one round of the task with high stakes. The median choice in the high-stakes decisions by the observers was 37 (the optimal choice), whereas the median choice by the subjects who engaged in the small-stakes task was 50. Merlo and Schotter offered this as evidence that the observers learned more effectively than the subjects engaged in the task.

Anderson and Holt (1997) studied an interesting situation in which equal weighting of personal information and observational learning (the information obtained by observing others' actions) leads to an information cascade (that is, it stops the accumulation of knowledge as decision makers stop using their private information). Their results show a lower rate of information cascade than predicted under the rationality assumption. This pattern can be explained by the assumption that people overweight their personal information. Clear support for this assumption is provided by Simonsohn et al. (2008). The participants in their studies received feedback concerning their payoffs (personal experience) and the payoffs of other agents. The results show that the effect of the personal experience was much larger than the effect of others' experience. Alos-Ferrer and Schlag (2009) review theoretical research that focuses on the value of imitation as a learning strategy. Their analysis demonstrates that payoffs affect the social value of imitation: Efficiency can increase by a tendency to rely on personal information if the advantage of imitation is small.

4 SOCIAL INTERACTIONS AND LEARNING IN GAMES

It is constructive to distinguish between two main effects of the social environment on choice behavior. First, the social environment can affect the strategies considered by the decision makers and/or the utility from the obtained payoffs. For example, it can lead the decision makers to consider strategies that facilitate reciprocation, increase fairness, and/or build trust. The second effect is indirect: the social interaction affects the obtained payoffs, and these payoffs shape behavior.

Most previous experimental studies of social interactions (games) focus on the direct "reciprocation-related" effects of the social environment (see Cooper and Kagel Chapter 4). The current review tries to complement this research by focusing on the indirect effect of the social environment. It builds on the observation (Roth and Erev

1995; Erev and Roth 1998) that there is wide set of situations in which the understanding of the obtained payoffs is sufficient to predict the outcome of social interactions. The effect of experience in this space of social situations is similar to the effect of experience in individual choice tasks, and it can be approximated with simple reinforcement learning models like I-SAW. One class of social interactions that belongs to this "basic shaping" space is the class of market-entry games examined in the choice-prediction competition described in Section 1.3.2. The best prediction of the outcome of this class of social interactions was provided by models that capture the basic properties of learning described in Section 1.1.

The main goal of the current section is to clarify the boundaries of the basic shaping space. Specifically, it examines the conditions under which the outcome of complex social interactions can be reliably predicted based on simple models that assume learning among the possible alternatives. In addition, it tries to shed light on the assumptions that have to be added to the basic models in order to capture behavior beyond this basic space.

Section 4.1 considers studies of learning in games under limited prior information. The results reveal examples of "emerged sophistication" that can be predicted with I-SAW and similar models.

Section 4.2 reviews studies of learning in 2-person constant sum games with unique mixed strategy equilibrium. The results reveal that prior information can affect the sequential dependencies in the data but has little effect on the aggregate choice rates.

Section 4.3 summarizes studies of cooperation and coordination. The results reveal that under certain conditions, players can learn to maximize efficiency by reciprocating and coordinating. In addition, the results suggest that this "learning-to-reciprocate" phenomenon is rather delicate. It is likely to emerge only when all the following six conditions are met: (1) the agents receive a reliable and complete description of the incentive structure, (2) the benefit from reciprocation is large, (3) the number of interacting agents is small (4 can be too large), (4) the noise level is low, (5) the interaction is expected to continue with high probability, and (6) the framing of the task clarifies the value of reciprocation. These results can be captured with the assertion that players consider "try-to-reciprocate" cognitive strategies. Yet the set of situations under which these strategies can be learned is not large.

Section 4.4 discusses studies that explore the role of fairness. The results show that in certain settings people behave as if they try to maximize fairness. However, in other settings they choose actions that reduce equality, even when this action impairs expected return. This pattern can be captured as another indication of considering, but not always using, try-to-reciprocate cognitive strategies.

Section 4.5 summarizes the main results and discusses alternative explanations and several open questions.

4.1 Social Interactions Given Limited Prior Information

4.1.1 The Group Size Effect in Mutual Fate Control Games

Sidowski, Wykoff, and Tabory (1956; and see Colman 2005; Colman et al. 2010; Delepoulle, Preux, and Darcheville 2000, 2001; Mitropoulos 2001, 2003) studied a minimalistic 2-person social situation in which the players can help each other but cannot affect their own payoffs directly. The top of Figure 10.7 presents a member of this class of "mutual fate" games that was studied in a 200-trial experiment by Colman et al. (2010).

	C	D
C	1, 1	0, 1
D	1, 0	0, 0

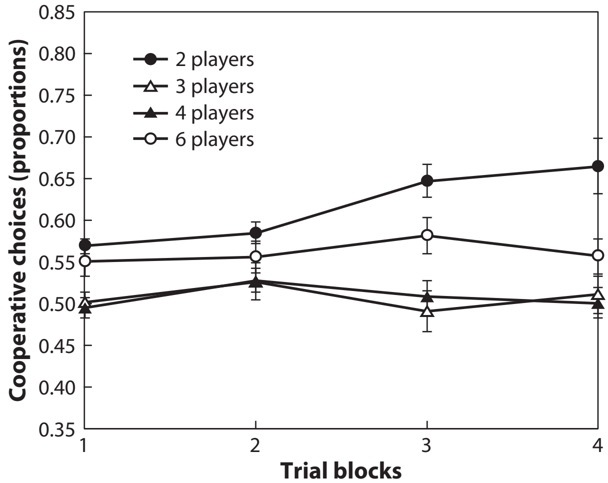

Figure 10.7: Mutual fate control game (*top*) and experimental results (*bottom*): proportions of cooperative choices over four trial blocks in groups of varying sizes. Error bars represent standard errors.

Notice that traditional game-theoretic analysis does not have clear predictions for the current game. Specifically, all 4 cells are weak Nash equilibria points in a 1-shot play of the game.[12] The participants in the typical experimental study of this class of games do not receive any information concerning the payoff rule and interact repeatedly in fixed pairs. The results show that most pairs slowly learn to coordinate on the efficient outcome (the “1, 1” cell). The proportion of efficient coordination after 100 trials is close to 70%.

Thibaut and Kelley (1959) show that this learning process can be a product of a *win-stay-lose-shift* (WSLS) decision rule. This rule implies a repetition of the last choice after high payoff and a change after a low payoff.

Colman et al. (2010) examine the effect of the number of interacting players in a multiplayer generalization of the mutual fate game. In the generalized game the players are placed in a ring, and each player has a predecessor on his or her left and a successor on his or her right. The payoff of each player is determined by her predecessor (the player receives 1 only if his or her predecessor chose C), and the action of each player determines the successor’s payoff.

The WSLC rule implies efficient coordination in multiplayer mutual fate games when the number of interacting agents is even (see Coleman, Colman, and Thomas 1990). Colman et al. (2010) experimental results, presented in Figure 10.7, do not support this prediction. Rather, they reflect a large qualitative difference between the basic $N = 2$ condition and the $N > 2$ conditions. The players learned to coordinate when $N = 2$ but not when $N > 2$. A similar group-size effect was documented by Feltovich, Iwasaki, and Oda (2007) in a study of a stag hunt coordination game.

Colman et al. (2010) show that this group-size effect can be captured with models that imply a stochastic WSLC decision rule and note that this class of models includes the leading model of decisions from experience in individual choice tasks (like I-SAW) presented in Section 1.

4.1.2 Quick and Slow Learning in Market-Entry Games

Erev and Rapoport (1998) document surprisingly fast convergence to Nash equilibrium in 12-person market-entry games that were played without prior information of the payoff rules. In each trial of one of these games participants chose between entering and staying out of the market. Staying out paid a sure payoff of 1. The payoff for entering was $1 + 2(4 - E)$, where E is the total number of entrants.

This game has multiple pure-strategy equilibria and one symmetric mixed-strategy equilibrium. The average number of entrants at these equilibria is between 3 and 4. The observed number of entrants in trials 15 to 20 (the last block) was 4.1, and the mean obtained payoff was between the expected payoff under the mixed and the pure equilibrium points.

At first glance, this coordination appears to contradict the low predictive value of the equilibrium predictions in the market-entry-game competition described in Section 1.3 (the ENO of the equilibrium prediction in this study was below 1). However, there is a simple explanation for the difference between the two studies. Erev and Rapoport examined situations in which the equilibrium prediction implies relatively small differences between the entry rate and the probability that entry is ex-post optimal (the proportion of trials in which entry leads to the best possible outcome). In these situations, learning toward equilibrium is relatively quick. The market entry game competition considered a wide set of games that includes cases with large differences between the equilibrium entry rate and the probability that entry is ex-post optimal. The results reveal that when this difference is large, learning toward equilibrium is slow, and the deviation from equilibrium can be described as reflection of underweighting of rare events.

4.2 Learning in Constant-Sum Games with Unique Mixed-Strategy Equilibrium

A two-person constant-sum game is a simplified social interaction that captures pure conflict: the sum of the payoffs of the two players is fixed, and the players cannot reciprocate. The game presented in Figure 10.8 is an example of a constant-sum game with a unique mixed-strategy equilibrium. In this equilibrium player 1 selects A1 with probability $p = 3/8$ and player 2 selects A2 with probability $\frac{7}{8}$. Under this mixed strategy, player 2 is expected to receive the same payoff from A2 ($\text{EV} = 0.7(\frac{3}{8}) + 0.6(\frac{5}{8})$) and from B2 ($\text{EV} = 0.2(\frac{3}{8}) + 0.9(\frac{5}{8})$). Thus, player 2 is not motivated to deviate from his predicted behavior. Similar logic holds for player 1.

4.2.1 Slow Learning and Limited Effect of Prior Information

Suppes and Atkinson (1960) examined the game of Figure 10.8 in a 210-trial experiment. The participants were run in fixed pairs: One participant was assigned to be player 1, and the second participant was assigned to be player 2. The payoffs are the winning probabilities. For example, if player 1 selects A1 and player 2 selects A2, then player 1 wins with probability 0.7 and player 2 wins with probability 0.3.

	A2	B2
A1	0.3, 0.7	0.8, 0.2
B1	0.4, 0.6	0.1, 0.9

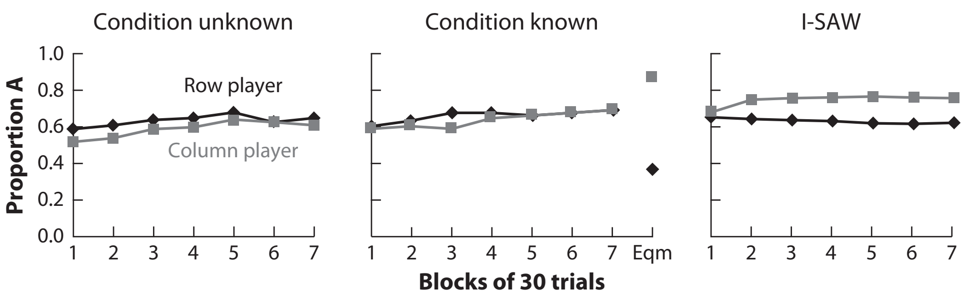

Figure 10.8: A constant-sum game study (from Suppes and Atkinson 1960) that reveals deviation from equilibrium, slow learning, and limited sensitivity to prior information.

Two information conditions were compared. The payoff matrix was known to the participants in Condition Known and unknown in Condition Unknown. The feedback after each trial was limited, in both conditions, to the realized outcome (Win or Loss).

The results, presented at the top of Figure 10.8, reveal a very small difference between the two conditions. The following observations summarize the results under both conditions: (1) The initial choice rates are close to 50%. (2) With experience player 2 increases the tendency to select A2. That is, Player 2 moves toward the equilibrium prediction. However, this movement is very slow. Even after 200 trials the proportion of A2 choices is closer to 50% than to the equilibrium prediction ($\frac{7}{8} = 87.5\%$). (3) Player 1 moves away from the equilibrium prediction: The observed proportion of A1 choices was above 60% (in equilibrium, player 1 is expected to select A1 in only 37.5% of the trials).

Follow-up research shows the robustness of the pattern documented by Suppes and Atkinson (1960). Slow learning and learning away by one of the players are quite common in constant-sum games with unique mixed-strategy equilibria. Ochs (1995) shows that a similar pattern can be observed in nonconstant-sum games that are played “against a population.” (The experiment was run in cohorts of 8 or more subjects in each role. In each trial all the participants in the role of player 1 played against all the participants in the role of player 2).

Erev and Roth (1998; see a clarification in Sarin and Vahid 1999) demonstrate that learning away by one player is predicted by simple models that assume exploitation (selection of the alternative that led to the best outcome in the past) and exploration/error (random choice). I-SAW is an example of this class of models. The right-hand column in Figure 10.8 shows the predictions of I-SAW (with the parameters estimated before) for the current game.

Additional indications of the robustness of these results are presented in Table 10.3. This table summarizes the results of experimental studies of three randomly selected constant-sum games. The games were run under two conditions. In Condition Minimal

TABLE 10.3:
Three of the randomly selected constant-sum games studied by Erev et al. (2002, 2007). The right-hand columns present the equilibrium prediction, the observed results, and the predictions of I-SAW.

Game (Probability that Player 1 Wins in Each Cell)						*Choice Rate Over 500 Trials by Information Condition*		
				Statistic	*Eq.*	*Minimal*	*Full*	*I-SAW*
		A2	B2					
1	A1	0.77	0.35	P(A1)	0.49	0.68	0.59	0.64
	B1	0.08	0.48	P(A2)	0.16	0.42	0.32	0.28
2	A1	0.73	0.74	P(A1)	0.99	0.76	0.84	0.84
	B1	0.87	0.20	P(A2)	0.79	0.40	0.36	0.21
3	A1	0.40	0.76	P(A1)	0.65	0.58	0.56	0.61
	B1	0.91	0.23	P(A2)	0.51	0.45	0.45	0.46

(see Erev et al. 2002), the participants did not receive a description of the payoff matrix, and the feedback was limited to the obtained payoff. In Condition Complete (see Erev et al. 2007) the participants received a complete description of the payoff matrix and complete feedback. Each game was run for 500 trials under fixed matching. The results show relatively small difference between the two information conditions (the correlation is 0.9), and learning away by one of the players in about half of the games. In addition, the results replicate previous studies (e.g., O'Neill, 1987) that demonstrate a relatively good match between the equilibrium predictions and the observed choice rate when the equilibrium predicts relatively small differences between the choice rates and the proportion of time that the different actions lead to the best payoff. In the context of 2×2 games, this condition holds when the equilibrium predictions are close to 50% (e.g., game 3 in Table 10.3).

The right-hand column in Table 10.3 presents the predictions of I-SAW (without reestimating the parameters—that is, based on the parameters used in Section 1 to fit the individual choice data) for the complete feedback condition. The MSD score is 0.0047 and the correlation is 0.93. This fit is better than the fit of the best model proposed in the original papers.

4.2.2 SEQUENTIAL DEPENDENCIES

It is important to recall that the quick learning toward the mixed-strategy equilibrium predictions, documented when the difference between the predicted choice rate and the implied success rate is relatively small (e.g., game 3 in Table 10.3), does not imply convergence to equilibrium. Studies of games in which aggregate behavior moves toward the equilibrium reveal that the sequential dependencies in the data differ from the predicted dependencies: Brown and Rosenthal (1990) reanalyzed O'Neill's (1987) results and found strong evidence of serial correlation in players' choices that contradict the equilibrium prediction (that imply no sequential correlations). The typical subjects exhibit overalternation. A similar overalternation bias was also documented by Rapoport and Budescu (1997) in a symmetric 2×2 game. Shachat (2002) shows that this deviation from the equilibrium emerges even when the players are allowed to use a randomization device.

Additional research suggests that the exact nature of the sequential dependencies in constant sum games is situation specific. For example, evaluation of the sequential dependencies in the constant sum games presented in Table 10.3 reveals that most subjects exhibit the opposite bias: Strong inertia (see Erev et al. 2007). Under one explanation of this pattern, overalternation emerges when the players are informed that they are selecting between objectively identical alternatives.

4.2.3 Modeling Robust Choice Rates and Slippery Sequential Dependencies

The constant-sum results presented earlier appear to reflect an interesting inconsistency: Section 3.2.1 suggests that the aggregate choice rates can be predicted with the assumption that behavior in different constant-sum games is driven by a general learning model like I-SAW, and Section 3.2.2 suggests situation-specific sequential dependencies. One resolution of this apparent inconsistency is based on the assumption that the different sequential dependency patterns are reflections of different situations- and person-specific exploration patterns that have limited effect on the aggregate choice rate (see a similar idea in Rapoport et al. 1997). This resolution can be naturally incorporated in a variant of I-SAW that allows for the possibility that during exploration, the agents tend to alternate between alternatives that are known to be similar.

4.3 Cooperation, Coordination, and Reciprocation

Rapoport, Guyer, and Gordon (1976) show that under certain conditions people can learn to cooperate in public good games and can learn to achieve efficient coordination. A clear demonstration of the emergence of cooperation is provided by the study of the prisoner's dilemma game presented in Figure 10.9 (Game PD1).

Each player in this 2-person normal-form game has to select between cooperation (C) and defection (D). When the game is played once, D is a dominant strategy (and the unique Nash equilibrium of the game). That is, each player earns more from selecting D than from C, independently of the choice of the other player. Yet, both players earn less when both select D (payoff of -1) than when they select C (payoff of 1).

In one of the experimental conditions, the participants played game PD1 for 300 trials against the same opponent (fixed matching) with immediate feedback after each trial (and without knowing how many trials would be played). The results (upper panel in Figure 10.9) show an increase in cooperation with experience. The cooperation rate in the last block was higher than 60%.

A clear indication of the emergence of coordination is provided by Rapoport, Guyer, and Gordon's (1976) study of following chicken game:

Game: Chicken 1	Swerve	Drive
Swerve	1, 1	−1, 10
Drive	10, −1	−10, −10

Notice that the game has two pure-strategy equilibria and one mixed-strategy equilibrium. The pure-strategy equilibria (swerve, drive and drive, swerve) are efficient (joint payoffs of 9) but unfair (one player wins 10 and the other loses 1). At the

Game (payoff matrix)

PD1	C	D
C	1, 1	−10, 10
D	10, −10	−1, −1

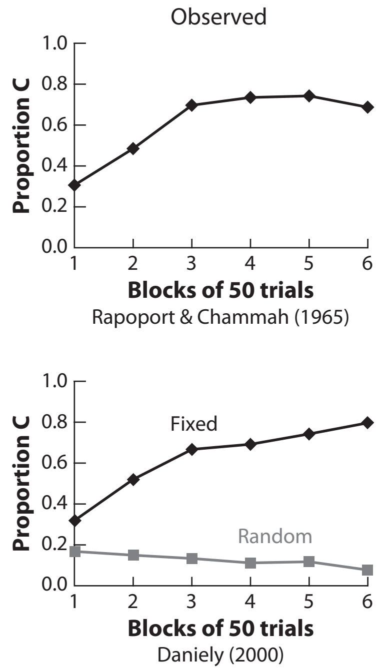

Figure 10.9: Two studies of a prisoner's dilemma game. The results reveal an increase in cooperation over time with fixed matching and a decrease with random matching.

symmetric mixed-strategy equilibrium, both players drive with probability $\frac{1}{2}$, and the expected payoff is 0. The results reveal that participants were able to achieve a high level of cooperation through alternating between plays of the game as to which player would drive and which would swerve. The efficient outcome (joint payoff of 9) was obtained in 84% of the trials. In addition, the results reveal a high level of fairness. The difference between the proportions of driving choices was lower than 7% for all 10 pairs.

Alternating behavior that facilitates efficiency and fairness was also shown by Arifovic, McKelvey, and Pevnitskaya (2006). They show that subjects playing repeated Battle of the Sexes, where there are two pure-strategy Nash equilibria, each favoring one player, often fall into a stable pattern of alternation between the two pure strategies. The data provided on the Web site accompanying the article shows that, out of 16 subjects matched in 8 fixed pairs, 56% individually alternated beginning in period 2. This is the proportion of subjects who chose a different action in period 2 than in period 1. By period 3, this proportion rose to 88% and by period 6, it reached 94%, which is all but one of the 16 subjects.

The emergence of cooperation and alternation-based coordination described here cannot be captured with basic reinforcement learning models like I-SAW. In the current context, human agents exhibit higher social intelligence and/or sensitivity than assumed by the basic learning models. In order to clarify the implications of this observation, the following sections review studies that highlight the conditions that facilitate sophisticated cooperation and coordination

The effect of the relative benefit from reciprocation. Rapoport and Chammah (1965) compare game PD1 with six other prisoner's dilemma games (same qualitative relationship between the different payoffs). Their results reveal high sensitivity to the relative benefit from cooperation. For example, when the benefit from unilateral defection was increased from 10 to 50, the cooperation rate decreased to 27%.

Size matters. The increase in cooperation with experience, discussed here, tends to weaken and even disappear as the number of interacting subjects gets large (e.g., Isaac and Walker 1988; Andreoni and Miller 1993; Daniely 2000; Huck, Normann, and Oechssler 2003; Bereby-Meyer and Roth 2006; Apesteguia 2006). That is, the likelihood of learning to cooperate is highly sensitive to the number of interacting agents. An increase in the number of interacting agents tends to increase the tendency to select the dominant strategy. A similar pattern was documented in the study of coordination games (Van Huyck et al. 1990; Bornstein, Budescu, and Zamir 1997).

For example, Daniely (2000) compared two versions of Rapoport and Chammah's prisoner's dilemma experiment (using game PD1). The first, referred to as "fixed matching," was a computerized replication of the original study. The participants were run in cohorts of 4 that were divided into 2 pairs. Each pair interacted 300 times. The results of this condition were very similar to the original results. The proportion of cooperation in the last block of 50 trials was 80%. The second condition, referred to as "random matching" (cf. Figure 10.9), was identical to the first, with the exception that the 4 participants in each cohort were randomly rematched after each trial. That is, the set of interacting agents over the 300 trials was increased from 2 to 4 (but the set of interacting agents in each trial was only 2 in both conditions). This change had a dramatic effect on the results. The proportion of cooperation in the last block of 50 trials dropped to 10%.

Apesteguia (2006) examined a 6-person public good game with and without description of the payoff rule. The results reveal very similar pattern in the two conditions. Another source of support to the suggestion that reciprocation is highly sensitive to the increase from 2 to 4 players is provided by Isaac and Walker (1988). They examined public good games (which can be described as generalized multiplayer prisoner's dilemma games). Their results showed a low cooperation rate in 4-player groups and similar rates with 7 agents (when the cost of cooperation is fixed).

Isaac, Walker, and Williams (1994) highlight an interesting boundary condition to the negative effect of group size on cooperation. Their results show that when an increase in group size increases the probability of very high payoffs from cooperation, it can eliminate the typical decrease in cooperation over time.

The role of framing. In addition to the two conditions described previously, Daniely (2000) studied the effect of framing. She tried to replicate the fixed matching study of game PD1 with the framing of the task as a transportation problem. Each player controlled a simulated car that approached a traffic light and had to decide between staying in his or her lane and changing lanes. The decision to change lanes increased the player's payoff and decreased the payoff of the other player. The exact payoff rule was determined by game PD1, with change implying D and stay implying C. As in the original study, the participant received a complete description of the payoff rule, and the feedback after each trial was complete. The only change between the studies was the addition of the transportation cover story. The results reveal that this addition eliminated the increase in cooperation. The observed cooperation rate in the last block of 50 trials was only 18%. Additional indications for the significance of the framing effect in the context of social interactions are presented by Rottenstreich (1995).

The shadow of the future. Selten and Stoecker (1986) studied behavior in a sequence of prisoner's dilemma games. Each player played 25 supergames, where each supergame consisted of a 10-round play of game PD2 (first panel in Table 10.4). Following each

TABLE 10.4:
Prisoner dilemma games that were studied using variants of Selten and Stocker's supergame procedure.

		C	D
PD2: Selten and Stoeker		C	D
	C	60, 60	−50, 145
	D	145, −50	10, 10
PD3: Andreoni and Miller		C	D
	C	7, 7	0, 12
	D	12, 0	4, 4
PD4: Bereby-Meyer and Roth		C	D
	C	0.105, 0.105	0.005, 0.175
	D	0.175, 0.005	0.075, 0.075
PD5: Dal Bó and Fréchette		C	D
	C	R, R	12, 50
	D	50, 12	25, 25

supergame, each player was rematched to a new opponent. The typical outcome was initial periods of mutual cooperation, followed by an initial defection, followed by noncooperation in the remaining periods. That is, the understanding that the game is about to end—or the lack of shadow cast by the future (Dal Bó 2005[13])—decreases endgame cooperation with experience. While early game cooperation increases with experience, so does endgame defection. Moreover, the first period of defection occurs earlier and earlier in subsequent supergames. Selten and Stoecker note that this learning pattern can be captured with a simple direction-learning model.

Andreoni and Miller (1993) studied game PD3 (second panel in Table 10.4) using the Selten and Stoeker prisoner's dilemma design. Their results replicated both the increase in early-round cooperation and the increase in late-game defection with experience between supergames documented by Selten and Stoeker. However, unlike Selten and Stoeker's finding that the defection period occurs earlier with experience, they find that the defection period occurs later with experience. The difference between the two studies could be attributed to the weaker temptation to defect in Andreoni and Miller's matrix. This interpretation of the results is consistent with findings by Dal Bó and Fréchette (2011) in the infinitely repeated PD games. Kagel and McGee (forthcoming), who studied finitely repeated PD supergames under the same paradigm—with individuals as well as teams—found that one factor determining whether subjects will defect earlier or later is the behavior of the partner. When the partner defected first in the previous supergame, subjects tended to defect earlier in the subsequent supergame. In essence, subjects are reacting to the past.

Noise matters. Bereby-Meyer and Roth (2006) examined the effect of payoff variability on choice behavior in a prisoner's dilemma game under Selten and Stoeker's supergame paradigm and under random matching. They focused on game PD4 (lower panel in Table 10.4). In the stochastic condition, the matrix entries represent the probability of winning $1. In the deterministic condition, the entries represent payoffs in cents. The results reveal an interesting interaction. Payoff variability increased cooperation given random matching but impaired cooperation under repeated play.

The effect of prior information. Coordination and reciprocation becomes very difficult when the agents do not know the incentive structure. As noted in Section 4.1, when the information is limited, coordination is difficult even in a common-interest game (Colman et al. 2010).

4.3.1 Alternative Abstractions: Social Utilities and Cognitive Strategies

Previous research highlights the value of the two main approaches capturing the effect of experience on cooperation and coordination. One approach is based on the importance of social utilities. For example, an increase in reciprocation can be captured with the assumption that successful reciprocation is reinforcing (see Macy and Flache 2002. Vega-Redondo 1997; Juvina et al. 2013). One recent demonstration of the potential value of this approach is the observation that people behave as if they find the act of following advice reinforcing (see Biele et al. 2009).

A second approach involves the assertion, discussed earlier, that people learn among a subset of repeated game strategies. For example, Erev and Roth (2001) assume that the player considers a "reciprocation" strategy that requires effort to reach the most efficient outcome (and punishes opponents who deviate from this play). When this strategy leads to good outcomes, players learn to select it. In another model (Hanaki et al. 2005), the players are assumed to consider strategies that can be represented by automata having no more than 2 states. Analysis of this model shows that it can capture the emergence of reciprocation. Alternative abstractions of the cognitive strategies idea involves a distinction between learning and teaching (see Camerer, Ho, and Chong 2002; Ehrblatt et al. 2006). Cooperation emerges under these models when sophisticated players are able to teach their opponents that cooperation is beneficial.

4.4 Fairness and Inequity Aversion

Studies of decisions from description demonstrate that in certain cases, people try to avoid inequity (increase fairness) even when this effort decreases their payoff (see the review in Cooper and Kagel in Chapter 4). Evaluation of the effect of inequity on learning reveals mixed results: Some studies show strong evidence for inequity aversion, but some studies suggest inequity seeking.

One demonstration of the effect of equity on learning is provided by Rapoport et al. in the prisoner's dilemma game described in Figure 10.9. Their results show almost perfect correlation between the payoffs of the two agents in each pair.

Another indication for inequity aversion is provided by studies of repeated ultimatum games (Guth, Schmittberger, and Schwarze 1982). In the basic version of this game, one player—the proposer—proposes a division of a pie (e.g., $10 in the experiment considered shortly) between himself or herself and a second player. In the second stage the second player—the responder—can accept or reject the proposal. If he or she accepts, each player gets the proposed share. If he or she rejects, both get nothing. The game-theoretic solution (subgame perfect equilibrium) states that the proposer should offer the smallest possible amount to the receiver, and the receiver should accept it. Abbink et al. (2001) examined a variant of this game in which the proposer's payoff, in the case of a rejection, was either 0 (as in the original game) or 10. Only the responders were informed of the proposer's payoff in the case of rejection. Responders were three times more likely to reject the unequal split when doing so enhanced equity (both players got 0) than when it reduced equity (when the rejection payoff to the proposer was 10 and 0 for the responder).

Indication for inequity seeking is provided by a study of betting games (Sonsino, Erev, and Gilat 2002; Erev et al. 2015). For example, in one of the conditions in

Erev et al., two agents have to decide between a participation in a zero sum bet and a safe prospect that implies a fair and efficient outcome (both agents gain 6 units). If both select the bet, one of them pays x units ($x = 10$, 20, 30 or 40) to the other agent. The game structure implied that rational consideration (and risk aversion and loss aversion) should lead the subjects to prefer the safe, efficient, and fair outcome. Yet the results reveal a high initial betting rate (about 80%) and very slow learning to stop betting. The betting rate after 250 trials with immediate feedback was around 50%. These results can be explained as a reflection of 2 of the regularities discussed earlier: The initial deviation from the fair equilibrium suggests that sensitivity to framing can be more important than inequality aversion, and the slow learning demonstrates the significance of the payoff variability effect.

4.5 *Summary and Alternative Approaches*

The current review of learning in social interactions shows three factors at play. First, learning in games can result in the emergence of reciprocation: in certain situations agents learn to increase their payoff by cooperating and coordinating. Second, the emergence of reciprocation can be captured with the assertion that the agents consider "try-to-reciprocate" cognitive strategies. Strategies of this type drive choice behavior when they are reinforced. Finally, the results suggest that there are many situations in which the effort to reciprocate has little effect on choice behavior. In these cases the effect of the incentive structure can be captured with the basic learning models presented in Section 1.[14]

5 APPLICATIONS AND THE ECONOMICS OF SMALL DECISIONS

The experimental studies reviewed earlier focus on small decisions: The stakes in the typical experimental task were small, and the participants did not invest very much time and/or effort in each choice. Nevertheless, we believe that the behavioral regularities documented in this research can be of high practical value. Our belief is based on three sets of observations. First, many important economic phenomena are the direct product of small decisions. For example, small decisions by drivers (e.g., the choice between the gas pedal and the brake pedal) affect traffic accidents, traffic jams, and pollution. Similarly, small clicking decisions by Internet users determine the future of newspapers and of the music industry.

Second, in many settings high-stakes decision problems are shaped by small decisions. For example, consider the high-stake decision among different job offers. In many cases this big decision problem is affected by earlier small decisions. The job offers available to a specific college graduate are likely to depend on small decisions that he or she has made as a child and as a student. Wise small decisions can help obtain high grades and build good social network that increase the probability of good job offers.

A third set of observations comes from studies that directly examine and demonstrate the practical implications of the learning phenomena reported on here. Some of these studies are reviewed next.

5.1 *The Negative Effect of Punishments*

The most influential contribution of the experimental analysis of learning is probably Skinner's (1953) clarification of the negative effects of punishment. Skinner focused on an environment in which (benevolent) "teachers" can use both reinforcements and punishments to shape the behavior of "students." His analysis shows that the overall effect of punishments can be negative even when they appear to be effective in reducing the frequency of the punished behavior. Specifically, the overall effect of punishments depends on the existence of "avoidance options": behaviors that differ from the shaping goals but can protect the students from punishments. An extreme example is the effect of punishments to facilitate effective reading and writing. When the teacher punishes errors, the student can learn to avoid these punishments by not coming to school.

Skinner's simple observation was among the most important triggers for policies that banned the use of corporal punishments in school. Analysis of the effect of these "less punishment" policies suggests that they are associated with a decrease in school dropouts and crime (Straus 1991).

Notice that Skinner's insight builds on three of the phenomena described before. First is melioration by the students. The tendency to avoid punishment by dropping out of school can be a reflection of insufficient sensitivity to delayed outcomes. A second phenomenon is the hot stove effect, which leads to convergence to a local maximum: Most students who had failed to master reading and writing could master these skills if they would have continued to explore different studying and remembering methods, but they gave up too early. Finally, the teachers' tendency to punish bad performance can be a reflection of underweighting of rare events (that can be the product of reliance on small samples). From the teacher's point of view, the common outcome of punishment tends to be positive (the students try harder), and the problematic avoidance reaction is rare.

5.2 The Enforcement of Safety Rules

The research reviewed in Sections 1– 4 has six implications for the design of safe working environments (see Erev and Rodensky 2004; Schurr, Rodensky, and Erev 2014; and related ideas in Zohar 1980). First, the results suggest that rule enforcement is necessary even when safe behavior (e.g., the use of safety equipment) is the rational course of action. The explanation of the relevant risks might not be enough. When workers make decisions from experience, they are likely to underweight the low-probability-high-hazard event and behave as if they believe it won't happen to them.

A second implication involves the negative effect of punishment, described earlier. Punishment can be an effective enforcement method only when the risk of problematic avoidance behavior is sufficiently low.

Two additional implications concern the effectiveness of rule-enforcement systems in which a small proportion of the violations are severely punished (see Becker 1968). The current review implies that systems of this type are likely to be effective in the context of decisions from description, but less effective, or ineffective, in the context of decisions from experience. When decisions are made from experience, low-probability punishments are likely to be underweighted. A related implication comes from studies of the importance of fairness considerations in social interactions, as in the ultimatum game experiments discussed earlier. This research suggests that the implementation of low-probability heavy punishments may be very difficult if the recipient can affect the enforcers (the proposer's role). Under the assumption that punishment may seem unfair (because only some violators are punished), some recipients are likely to retaliate even if retaliation is costly to them. Thus, enforcers (proposers) might learn to avoid using these punishments.

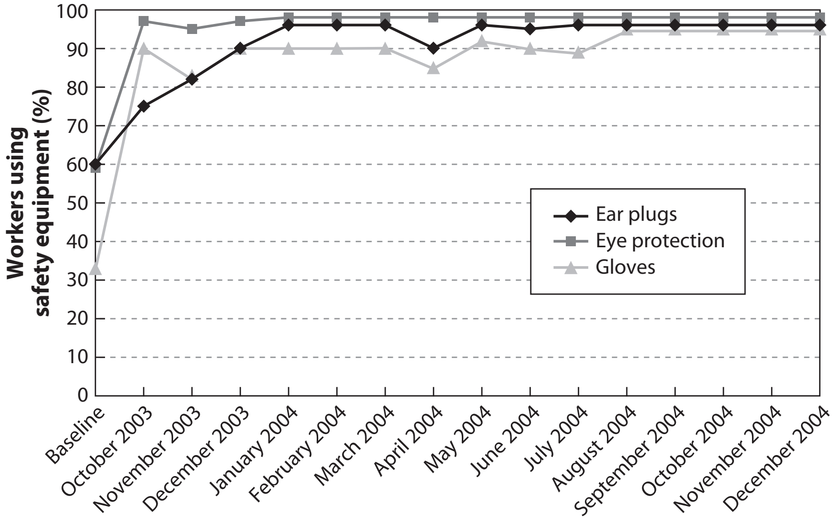

Figure 10.10: Percentage of workers who obey the safety rule and use the required safety equipment as a function of time in one of the departments (studied by Schurr, Rodensky, and Erev 2014).

A fifth implication is optimistic. It implies that the fact that workers take unnecessary risks and behave as if they ignore safety rules does not imply that they will object to attempts to enforce these rules. Indeed, the observation that low-probability events are overweighted in decisions from description implies that when workers are explicitly asked to consider the safety issue, they will agree that they want to behave safely and will be happy to see that the management designs a rule-enforcement system to help them achieve this goal.

Finally, the arguments just presented suggest that behavior is much more sensitive to the probability than to the magnitude of the punishment. Thus, a gentle continuous punishment ("gentle COP") policy that implies low punishments with high probability can be very effective (as long as the fine is larger than the benefit from violations of the rule and the risk of avoidance behavior is low).

Erev and Rodensky (2004; and see Erev 2007; Schurr, Rodensky, and Erev 2014) applied this gentle COP method in 12 Israeli factories. The basic idea was to design a mechanism by which supervisors will be encouraged to approach each worker who violates the safety rule and remind him or her 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 prior to intervention. The data included objective measures of the workers' safety behaviors (cf. Figure 10.10). The intervention started with a formal presentation of the new policy to all the workers. Figure 10.10 presents measures of safety related behavior before and after the presentation in one of the departments in one of the twelve factories. The baseline data were collected by the research team a month before the beginning of the intervention (in September 2003) and were independent of the supervisors' comments and records.

As demonstrated in Figure 10.10, the intervention had a large and immediate positive effect. A similar pattern was observed in all 12 factories. The rate of safe behavior increased to 90% immediately after the beginning of the intervention. More interesting is the observation that the effect of the intervention did not diminish with time. The rate of safe behavior increased or stayed high during the 2 years since the beginning of the intervention. Given the success of the intervention and its relatively low cost, the factories have decided to maintain the experimental policy.

5.3 *Cheating in Exams*

One of the likely contributors to the long-term success of the gentle COP procedure is the observation that multiple equilibria are common in rule-enforcement problems, including tax compliance (Alm and McKee 2004) and corruption (Tirole 1996; Waller, Verdier, and Gardner 2002). In one equilibrium, obeying the rules is the norm, and the enforcers 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. The possibility of two extreme equilibria and the hypothesis that small decisions are made based on experience in similar situations implies that the effectiveness of different rule-enforcement policies is likely to be particularly sensitive to the initial actions. Wise allocation of initial resources can lead to a convergence to the "good" equilibrium, in which observing the rule is the norm.

Erev, Ingram, et al. (2010) applied this reasoning to cheating on college exams. Their analysis suggests that gentle COP policies can be used to move behavior to the good equilibrium. To evaluate this hypothesis, they ran an experiment during final semester exams of undergraduate courses at the Technion. Traditionally, instructions for exam proctors at the Technion included the following points:

1. The student's ID should be collected at the beginning of the exam.
2. A map of students' seating should be prepared.[15]

Since 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 exam. Early preparation of the map reflects an attempt to follow Becker's idea (preparing evidence to facilitate large punishments) but distracts the proctors and reduces the probability of gentle punishment (e.g., warning or moving the suspected student to the first row) at the beginning of the exam.

The experiment compared two conditions that differed with respect to the timing of the preparation of the map. In the control condition, the proctors were requested to prepare the map at the beginning of the example (as they did before the study), and in the experimental condition, they were requested to delay the preparation of the map by 50 minutes.

Seven undergraduate courses were selected to participate in the study. In all courses the final exam 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 exam, students were asked to complete a brief questionnaire in which they are asked to rate the extent to which students cheated in this exam relative to other exams. The results reveal a large and consistent difference between the two conditions. The perceived level of cheating was lower in the experimental condition in all seven comparisons.

5.4 Broken Windows Theory, Quality of Life, and Safety Climate

In an influential paper, Kelling and Wilson (1982) suggest that physical decay and disorder in a neighborhood can increase the crime rate. This suggestion, known as the broken windows theory, was motivated by a field experiment conducted by Zimbardo (1969). The experiment focused on two cars that were abandoned in the Bronx, New York, and in Palo Alto, California. The results showed that vandalism of the cars started only after the experimenter created disorder (by removal of the license plate or breaking a window).

The broken windows theory was a motivation for the "quality-of-life" policing strategy implemented in New York City in the mid-1990s (Kelling and Sousa 2001). This policing strategy advocated increasing the number of police on the streets and arresting persons for less serious but more visible offenses. Some credit this strategy for the decline in crime and disorder (Golub et al. 2002; Kelling and Sousa 2001, Silverman 1999). However, there are other explanations for the decline (see Eck and Maguire 2000). Field studies that test the broken windows hypothesis provide mixed results. Skogan (1990) found that robbery victimization was higher in neighborhoods characterized by disorder, but Harcourt (2001) found that the crime-disorder relationship did not hold for other crimes, including burglary (housebreaking), assault, rape and pickpocketing.

We believe that the studies reviewed here can help clarify this mixed pattern. Under the current analysis, quality-of-life policing can be effective for the same reason that gentle COP policies are effective. When the probability of detection is very high and the risk of problematic avoidance behaviors is low, people learn to obey the rule. Thus, quality-of-life policing is effective in reducing robberies because these violations are more likely to be detected by the additional neighborhood police.

Luria, Zohar, and Erev (2008) examined this "probability of detection" explanation in the context of a safety-climate intervention (Zohar 1980). Safety-climate interventions are very similar to quality-of-life policing. These interventions are designed to create a safer work climate. This goal is achieved by encouraging supervisors to exhibit commitment to safety (e.g., by measuring the number of times they discuss safety issues with their subordinates). Zohar (1980) and Zohar and Luria (2005) show that this manipulation increases safety. To test the probability of the detection hypothesis, Luria et al. reanalyzed the data reported in Zohar and Luria. Their results show that the safety climate decreases unsafe behavior in environments with high visibility (the supervisor can detect rule violation with high probability) but not when visibility is low.

Notice that this explanation for the effect of quality-of-life policing has nontrivial positive and negative implications. On the positive side, this explanation implies that it may not be necessary to arrest all violators of minor crimes. If the probability of detection is high enough, more gentle punishment may be enough. For example, if the probability of detecting an attempt to use public transportation without paying is close to 1, then a fine that is only slightly larger than the regular cost should be sufficient. On the negative side, the current analysis suggests that quality-of-life policing is not likely to succeed when the probability of detection is low.

5.5 Hand Washing

Hand washing is a nice example of the difference between decisions from experience and decisions from description. The consequence of a failure to wash one's hands is potentially devastating—including serious illness or even death. The cost of washing one's hands is a few seconds of inconvenience. Everything we know about decisions from description—including risk aversion, loss aversion, and overweighting of small probabilities—suggests that people would be eager to wash their hands. Yet, repeated experience following not washing one's hands is likely to result in no noticeable negative outcome and, therefore, in extinction of this desirable behavior.

In 1847, Dr. Ignaz Semmelweis first demonstrated that routine hand washing could prevent the spread of disease. In an experiment, Dr. Semmelweis insisted that his students staffing a Vienna hospital's maternity ward wash their hands before treating the maternity patients—and deaths on the maternity ward fell dramatically. In one case, it fell from 15% to near 0%! Though his findings were published, there was no apparent increase in hand washing by doctors until the discoveries of Louis Pasteur years after Dr. Semmelweis died in a mental asylum (Nuland 2003).[16]

Moreover, many believe that even today medical professionals do not do enough on this front. In a recent study, Erev, Rodensky, et al. (2010) used a variant of the gentle COP policy, described earlier, to increase the use of gloves by doctors and nurses. They focused on the use of gloves while taking blood and giving infusions in 12 distinct departments. The gentle intervention consisted of a single meeting with the department staff. During this meeting the researchers suggested that the participants help each other remember to use gloves. That is, when they see a friend approach a patient without new gloves, they should ask him or her to fix the problem. The results show that this minimal manipulation increased glove use from 50% to 95%.

5.6 *The Effect of the Timing of Warning Signs*

Evaluation of the impact of warnings reveals a large effect of prior experience (see Barron, Leider, and Stack 2008). Individuals who have had good experiences in the past are less affected by warnings. For example, when the FDA added a black-box warning to the drug Cisapride, the data show an increase in usage of 2% among repeat users, but a decrease of 17% among first-time users (Smalley et. al. 2000). Another example is provided by a study of parent-adolescent sexual communication. Regular condom use was found to be lower when parent-adolescent sexual communication occurred at a later age (Hutchinson 2002) as students had presumably already engaged in unsafe sexual activity and found it pleasant. Barron et al. show that the effect of experience remains even after controlling for the available information. Indeed, experience reduces the tendency to respond to informative warnings even if the experience does not provide additional information. It seems that part of the effect of experience is to underweight warnings as a result of inertia.

5.7 *Safety Devices and the Buying-Using Gap*

The difference between decisions from experience and decisions from description suggests that in certain cases people may buy safety devices but "learn" not to take the necessary measures to benefit from them. One example of this buying-using gap is a study by Yechiam, Erev, and Barron (2006) that focuses on car radios with a detachable panel. The detachable radio panel was (around the end of 20th century) a rather popular example of a safety device (against theft) that can be effective only when it is used (detached).

Notice that the main role of a detachable panel to a car radio is its value as a safety device. The decision not to detach the panel is made without explicit presentation of a threat and is likely to be shaped by repeated experience. Thus, the properties of decisions

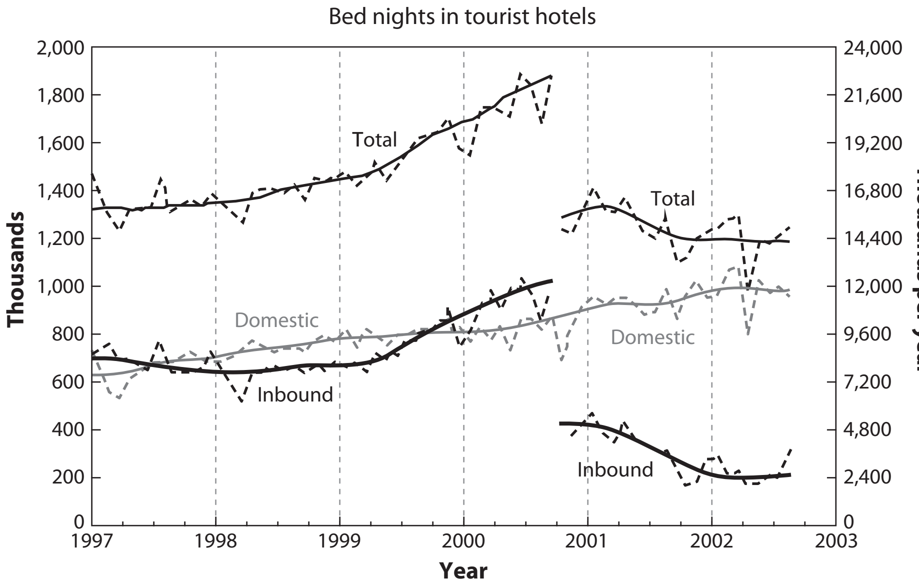

Figure 10.11: Bed nights in tourist hotels in Israel from January 1997 to August 2002: seasonally adjusted average (*dashed line*) and trend by 1,000 bed nights. Source: ICBS (2002b); used with permission.

from experience imply a decrease in the tendency to use the panel with experience, since the small probability of theft is underweighted. Yechiam et al. found (using a short survey) that the large majority (96%) of Israelis who bought car radios between 1995 and 2003 preferred the type with a removable panel even though it was more expensive. Most participants detached the panel in the first 2 weeks and were much less likely to detach it after a year. That is, responders behaved as if they gave more weight to the probability of theft in their initial-use decisions than in their use decisions after a year of experience.

5.8 The Effect of Rare Terrorist Attacks

Previous studies reveal that even rare terrorist attacks can have large negative effects on international tourism. For example, following terrorist activity in Northern Ireland in the early 1970s, visitors fell from close to 1 million in 1967 to about 300,000 in 1976.

Yechiam, Barron, and Erev (2005) note that the research just reviewed implies that other effects of terrorism may not be as large. Specifically, it implies a large difference between international and local tourism. Traveling to a different country requires a big decision from description. Local tourism, on the other hand, can be a product of small decisions from experience (e.g., whether to take a sandwich to work or dine in a restaurant) and can be affected by experience. Thus, with experience, the effect of rare terrorist attacks on local residents is likely to decrease.

Figure 10.11 presents the number of nights slept in Israeli hotels by local and international tourists before and after the beginning of the wave of terrorist attack that started at September 2000 and lasted several years. The results show a drop for both populations with the beginning of the recent attacks but a quick recovery by local tourists. This trend is consistent with the suggestion that experience reduces the impact of rare attacks.

Yechiam et al. note that their analysis suggests that the negative effects of rare terrorist attacks (on the economy) can be reduced by ensuring that citizens continue to partake in relatively safe leisure activities. Interestingly, this suggestion summarizes one component of Mayor Rudolph Giuliani's response to the September 11 attack in New York City. Mayor Giuliani suggested that citizens should invest less in direct contributions (like helping digging and collecting blankets), and spend more time shopping and dining in New York. While this suggestion seemed counterintuitive at the time, the current analysis suggests that it was effective in reducing the negative long-term economic effect of the attack.

5.9 *Emphasis-Change Training, Flight School, and Basketball*

Mane and Donchin (1989) have organized an interesting competition between leading researchers of motor-skills learning. The participants in the competition were asked to develop a training method to improve performance in a complex "space fortress" video game. The human players in this game control a spaceship and try to destroy a space fortress that tries to destroy their ship (using missiles and mines). High performance in this game requires sensitivity to several sources of information (e.g., the location of mines, the movement of missiles, the location of the ship, and the angle of the ship's gun).

One of the most successful submissions to this competition, proposed by Gopher Weil, and Siegel (1989), was based on the idea of *emphasis-change training*. During training, under this method, the trainees are continuously asked to change their focus. For example, they start by trying to maximize their scores on hitting the fortress, and then they are asked to focus on avoiding mines. The basic idea behind this method is simple: under the assumption that people choose among multiple attention-control strategies, they are likely to converge to a local maximum (see Section 1.3.2). Emphasis change reduces the risk of this problem (see Erev and Gopher 1998) by giving the trainee experience with attention-control strategies he or she might not otherwise sample.

The emphasis-change method was a clear winner in transfer tests (see Fabiani et al. 1989). One demonstration of the value of this method is provided by Gopher, Weil, and Bareket (1994). In the experimental group of their study, cadets in flight school were asked to play the space fortress game and practiced using the emphasis-change training method. The results reveal that this experience had a large positive effect on their subsequent performance in flight school. The probability of successful completion of the course increased by 33%.

Another demonstration of the value of emphasis-change training is provided by the success of a commercial variant of the space fortress game (see www.intelligym.com) designed to facilitate attention control by basketball players. The commercial product was used by only two NCAA men's basketball teams in 2005: the University of Memphis and the University of Florida. Florida won the NCAA title in both the 2005–6 and 2006–7 seasons. Twelve NCAA teams used the emphasis change trainer in the 2007–8 season: one of them (University of Kansas) won the title and another user (University of Memphis) was the runner-up.

5.10 *The Pat-on-the-Back Paradox*

Informal rewards, often referred to collectively as pats on the back, are low-cost or no-cost, often verbal, rewards that have virtually no monetary market value. Psychological research has shown that pats on the back can be as motivating as monetary awards. For example, Stajkovic and Luthans (1997) present a meta-analysis of 19 studies showing that feedback and social reinforcers may have as strong an impact on performance as monetary rewards. Survey-based data suggest similar conclusions. In a survey of American workers, 63% indicated a pat on the back to be an effective incentive (Lovio-George 1992). In other survey-based studies (Graham and Unruh 1990), pat-on-the-back incentives are shown to be more effective than monetary rewards. Such findings are often attributed to the recognition bestowed by the pat on the back and have prompted statements such as, "There are two things people want more than sex and money . . . recognition and praise" (Nelson 1994, quoting Mary Kay Ash, founder of Mary Kay Cosmetics).

These results appear to be inconsistent with the observation that most job postings focus on the salary, opportunities, and the possibility of promotion and professional development, and not on the likelihood of pats on the back. Luria et al. (2016) show that this "pat-on-the-back paradox" can be resolved as a reflection of the differential weighting on rare events in decisions from experience and from description. This explanation is based on the assumption that the probability of explicit monetary rewards (like promotions and bonuses) in the typical workplace is low. Thus, these events are overweighted when considering a description of the job but are underweighted in decisions from experience. Underweighting of rare rewards is expected to reduce effort in the workplace. To address this problem, wise managers use pats on the back as "lottery tickets" that signal a probabilistic future value (like a possible promotion), thereby reinforcing the behavior in question.

5.11 Gambling and the Medium-Prize Paradox

According to the leading explanations of gambling, people gamble because they overweight rare events (Kahneman and Tversky 1979) or because they are risk seekers around the status quo (Friedman and Savage 1948). These factors can explain the popularity of gambling games that promise positively skewed payoff distributions that provide very high payoffs with very low probability. However, they appear to be inconsistent with the observation that a large proportion of the payoffs in many gambling games involve medium prizes. Medium prizes are particularly common in casino settings.

Haruvy, Erev, and Sonsino (2001, following Skinner 1953) suggest that the coexistence of high and medium prizes can be a response to two behavioral biases: overweighting of rare events in decisions from description and the payoff variability effect in decisions from experience. High prizes are necessary to attract new gamblers (who respond to a description of the game), and medium prizes are necessary to increase the payoff variability that slows learning (that gambling is costly).

5.12 The Evolution of Social Groups

Recent research demonstrates that two of the most basic observations from studies of the development of social groups can be a product of the hot stove effect. Denrell (2005) focuses on the observation that proximity is an important determinant of liking (Brewer and Campbell 1976; Festinger, Schachter, and Back 1950; Segal 1974). Even if students are randomly assigned to rooms, individuals are more likely to become friends with and have a favorable impression of individuals who are nearby (Segal 1974). Denrell's explanation is simple and elegant: our opinions about our friends are likely to change after each meeting. When these opinions determine the probability of future meeting, we will stop meeting a friend when we no longer like him or her (and keep our low opinion). This problem is less severe when the proximity is high. For example, roommates meet independently of changes in their contemporary opinions. Thus, proximity limits the hot stove effect in this setting.

Denrell and Le Mens (2007) extend this analysis and show that the hot stove effect can partially explain why friends hold similar beliefs. This observation is based on the assumption that low evaluation of an activity (like eating at a particular restaurant, or attending service at a particular church) decreases the probability of a repetition of this activity. Friendship slows this process because high evaluation by a friend can lead us to repeat activities even when our personal evaluation is low.

Another example of a possible effect of decisions from experience to the development of social groups involves the survival of sects and religious groups that demand significant sacrifice. As noted by Berman (2001), successful groups appear to create an incentive structure in which the cost of exiting the group increases over time. Thus, melioration and related properties of decisions from experience can be among the contributors to the success of these groups.

5.13 *Product Updating*

Consumers have long been known to exhibit inertia in moving from one technology standard to another, even when the newer standard is demonstrably superior (Clements 2005; Gourville 2003). Microsoft, for example, the largest and most successful computer software company, is often criticized on the grounds that its products are inferior to competitors' products. Nevertheless, Microsoft products are often dominant in the market. While the reasons behind Microsoft's dominance are complicated and numerous (including the importance of establishing a network of users, complementarities, and unfair anticompetitive practices by Microsoft), research on consumption of other experience goods (products that require consumption before knowing their quality) has shown that consumers who behave as hill climbers will be unable to move easily from the old to the new product and will often converge to a local maximum.

Consumer learning in experience goods markets has been an important subject of theoretical research in industrial organization and marketing since the 1970s. Learning can be an especially important factor in the demand for new products, and there is an empirical literature that quantifies learning in household panel data for grocery purchases (for example, Erdem and Keane 1996), choice between personal computers (Erdem Keane, and Oncu 2005), and choice between drugs (Crawford and Shum 2005). In these papers, it is assumed that the only type of demand dynamics comes from learning, which creates inertia, partially explaining the reluctance of Microsoft consumers to switch to superior products. Likewise, this explains why many consumers do not immediately switch from a product they currently use to the latest improved product, even if the cost difference is minimal (Gourville 2003). He finds support for the basic learning assumptions described here: consumers are sensitive to relative payoffs of the two products, and their reference points about each product's quality critically depend on past experience. Local hill climbing can therefore take consumers to a suboptimal product choice and keep them there.

5.14 *Unemployment*

The decision to accept a particular job offer is often not a small decision. The stakes are usually high, and the decision maker is likely to invest time and effort in this choice.

TABLE 10.5:
The asymmetric stag hunt game considered by Erev and Greiner (2015).

	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

Nevertheless, many small decisions are likely to affect the employment status of the decision maker. Examples include the decisions to invest effort in particular tasks in school, at work, and while looking for a job. These small decisions are likely to affect the likelihood of receiving attractive job opportunities.

Lemieux and MacLeod (2000) present an elegant analysis that demonstrates how the basic properties of learning, reviewed earlier, can shed light on an apparently weak relationship between unemployment rates and public policies. They focus on the unemployment rate in Canada in the period of 1972 through 1992. The Canadian unemployment insurance system greatly increased benefits to the unemployed in 1971. The generosity of the unemployment insurance did not increase again, but unemployment steadily increased from 1972 to 1992. Lemieux and MacLeod note that this pattern can be captured with the assertion that the description of the incentive system has limited effect. The main effect is a result of personal experience with the new incentives.

5.15 *Interpersonal Conflicts and the Description-Experience Gap*

Review of research on interpersonal conflicts reveals an apparent inconsistency between the main conclusions of two major lines of research. On one hand, extensive research in behavioral game theory highlights the importance of other-regarding preferences (see Fehr and Schmidt 1999; Bolton and A Ockenfels 2000; Charness and Rabin 2002; and the review in Cooper and Kagel, Chapter 4). This research suggests that people pay more attention to the incentives of others than predicted under traditional assumptions of fully rational economic man. On the other hand, negotiation research reflects "mythical fixed pie beliefs" (see Bazerman and Neal 1992) that imply the opposite bias: a tendency to ignore the incentives of others and assume that efficient cooperation or coordination is impossible.

Erev and Greiner (2015) suggest that this apparent inconsistency can be a product of the difference between decisions from description and decisions from experience discussed earlier. It is possible that social behavior reflects oversensitivity to the outcomes of others when these outcomes are described (the convention in mainstream behavioral economic research) but reflects the basic properties of decisions from experience when the outcomes are not clearly described (the state in most negotiation settings). The basic properties of decisions from experience, in turn, imply a tendency to exhibit insufficient sensitivity to the payoffs of other agents.

Erev and Greiner clarify this assertion with the study of the 5×5 asymmetric stag hunt game presented in Table 10.5. Notice the game has two equilibrium points: The "E, E" equilibrium is efficient (payoff dominant) and fair: both players win 12 (joint payoff of 24) under this equilibrium. The "A, A" equilibrium is inefficient (joint payoff of 15) and unfair (one player wins 10, and the other wins 5), but it is the risk-dominant equilibrium. The game was played repeatedly (for 50 trials), with fixed matching, under 2 information conditions. The participants received a complete description of the matrix in the Description condition but not in the Experience condition. The results reveal a large difference between the two conditions. The modal outcome was efficient and fair ("E, E"—as predicted by other-regarding preferences) in Description condition, and inefficient and unfair ("A, A"—as predicted by the basic properties of decisions from experience) in the Experience condition.

The current analysis leads to optimistic predictions: It implies that manipulations that increase exploration (like the emphasis-change procedure described in Section 5.9) can increase social efficiency. This prediction is consistent with the main idea of popular negotiation books.

5.16 *Implications for Financial Decisions*

Typical financial decisions often involve high stakes. Nevertheless, recent research demonstrates interesting similarities between financial decisions and the experimental literature reviewed here.

The best-known example is provided by Taleb's (2007) prediction of the 2008 financial crisis. Taleb used the tendency to underweight rare events in decisions from experience, reviewed earlier, to justify his "black swan" assertion, according to which investors tend to dismiss low-probability events. For that reason, low-probability events, when they occur, can lead to financial crises.

Another example involves the assertion that many investors have underdiversified investment portfolios (e.g., Blume and Friend 1975; Kelly 1995). Ben Zion et al. (2010) show that this tendency can be observed in the clicking paradigm and can be a product of the tendency to rely on past experience.

A third example concerns sequential dependencies in stock markets. Empirical analyses reveal high correlation between absolute price change in a particular trading day and volume of trade in the following day (see Karpoff 1988). Nevo and Erev (2012) show that this pattern can be a product of the surprise-trigger-change of decisions from experience.

5.17 *Summary and the Innovations-Discoveries Gap*

The first author of the current chapter was recently invited to give a talk in a lecture series with the title "Inventions and discoveries that have shaped the human civilization." While preparing the talk, he noticed a surprisingly large gap between his favorite examples of inventions and discoveries in economics. Whereas the most influential inventions (e.g., markets, money, banks, rules, credit cards, auctions, e-trading, matching) are based on the assumptions that people try to maximize expected return, many of the interesting discoveries reflect deviations from maximization.[17]

We believe that the results reviewed here highlight one contributor to this gap. The basic properties of decision from experience imply interesting deviations from maximization but also imply a wide set of situations in which people behave as if they are trying to maximize expected return: When the strategy that maximizes expected return also leads to the best outcome most of the time, people exhibit a high sensitivity to the incentive structure. (This prediction is clarified by I-SAW: when the best alternative is also best most of the time, the "grand mean" and the "sample mean" tend to point in the same direction). It seems that many of the successful economic innovations are mechanisms that increase the probability that the socially desired behavior will be reinforced most of the time.

Most of the applications considered here follow a similar logic. They start with the discovery of a problematic deviation from maximization that can be the product of the tendency to rely on small samples and then show that the problem can be addressed by a change of the incentive structure that increases the probability that the desired behavior will be reinforced on average—and most of the time.

6 CONCLUSION

The research reviewed here can be summarized by six main sets of observations. The first set includes demonstrations of the generality of basic properties of decisions from experience. These behavioral regularities have been observed in animal studies, laboratory studies that focus on the behavior of student subjects engaging in simple tasks, and in the analysis of relatively complex social interactions. An additional indication of the robustness of the main results is provided by the observation that they can be summarized with a simple model (best reply to a small sample of experience in similar situations) that allows for useful ex ante quantitative predictions of behavior in new situations.

A second set of observations involves two shortcomings of an approach based on the strictest interpretation of rationality—including equilibrium analysis. First, there are many situations in which this approach leads to ambiguous conclusions (it is "not even wrong"). For example, this approach does not provide a clear prediction of behavior in the clicking paradigm. Almost any behavior can be justified as "rational," given certain prior beliefs. Second, when the rationality assumption leads to unambiguous predictions, it is often wrong at the intermediate term. For example, learning away from a mixed-strategy equilibria persists for at least 500 trials (see Section 4.2), and learning away from a simulated index fund that is known to maximize expected payoff and minimize variance experience persists for at least 100 trials (see Section 1.3.1). It is important to recall, however, that the current results do not reject the class of "epsilon equilibrium models" (e.g., Radner 1980; McKelvey and Palfrey 1995). Indeed, the descriptive models presented before are members of the class of epsilon equilibrium models: When the incentive structure is strong enough (in the way implied by these models), they imply an approximation of the optimal behavior.

A third set involves the conditions under which experience leads decision makers toward maximization of expected return (and risk-neutral equilibrium). High maximization rate was documented when the strategy that maximizes expected return also leads to the best outcome most of the time. Similarly, convergence to mixed-strategy equilibrium was observed when the choice proportions at equilibrium are consistent with the proportions of times in which each alternative leads to the best outcomes.

A fourth set of observations concerns the difference between decisions from experience and decisions from description. The results described here suggest that decision makers underweight rare events in decisions from experience but overweight rare events in decisions from description (see Section 1.1.3). Another example of this is the apparent inconsistency between research documenting other-regarding behavior and the finding that some social conflicts reveal the opposite bias (see Section 5.15).

The fifth set pertains to the distinction between basic learning properties and other cognitive factors that affect the impact of experience. The current review suggests that the effects of other cognitive factors are important but are less general than the basic properties of learning. For example, the indications for learning to follow a reciprocation strategy in a repeated prisoner's dilemma game are highly sensitive to the framing of the task.

Finally, the current review suggests that the study of decisions from experience may shed light on many interesting economic phenomena. Highly consequential economic phenomena may be the result of small and relatively inconsequential decisions by many individuals. The applications presented in Section 5 suggest that experimental research on small decisions can be used to understand larger phenomena and facilitate efficient design of relevant incentive structures.

ACKNOWLEDGMENTS

Much of this paper was written when Ido Erev was a Marvin Bower Fellow at Harvard Business School. It was also supported by the I-CORE program of the Planning and Budgeting Committee and the Israel Science Foundation (grant no. 1821/12). We thank Olivier Armantier, Greg Barron, Gary Bolton, Eyal Ert, Dan Friedman, Glenn Harrison, Teck Ho, John Kagel, Elena Katok, Steven M. Kemp, Amnon Rapoport, Al Roth, Andrew Schotter, Uri Simonsohn, Trent Smith, Dale Stahl, Nat Wilcox, and Eldad Yechiam for useful comments.

NOTES

1. "It is unnecessary to assume that the participants have full knowledge of the total structure of the game, or the ability and inclination to go through any complex reasoning process" (Nash 1950, 21)
2. Smith showed that competitive equilibrium could be attained with small numbers of buyers and sellers with no knowledge of others' costs or values.
3. Another reason for our interest in small decisions is the feeling that external validity of laboratory research is larger in the context of small decisions that are similar to the laboratory tasks in many ways (e.g., low stakes, limited decision time) than in the context of large decisions. So, we have more to say about small decisions.
4. Erev and Livene-Tarandach (2005) showed that standardized experimental paradigms could be used to reduce differences between natural sciences and social sciences. Many exam questions in the natural sciences (about 64% in the sample of physics GRE exams used to evaluate applicants to graduate school) and few questions in the social sciences (about 10% of the questions in psychology GRE exams) require predictions. The focus on standardized experimental paradigms could be used to reduce this gap by facilitating the development of short and clear prediction questions in the social sciences.
5. The payoff variability effect is related to the role of flat payoff functions. Harrison (1989) notes that the deviation from maximization (and equilibrium) observed in many experimental studies can be a product of the low expected cost of these deviations relative to the required effort to find the optimal choice. Merlo and Schotter (1992) refine this assertion and note that there may be large differences between the expected and the experienced costs. The payoff variability effect suggests that the best predictor of these deviations is the relative cost: the average cost relative the payoff variance. This suggestion is consistent with Harrison assertion under the assumption that payoff variability is one of the factors that increases the effort required to find the optimal choice.
6. The probability mixture (B, p) denotes a win prospect B with probability p and 0 otherwise.
7. Additional research suggests that the importance of rare events is best approximated by the difference in expected values relative to payoff variance.
8. In addition to this competition, Erev et al. (2010a) organized a competition that focused on decisions from description and a competition that focused on decisions based on free sampling. The comparison of the three competitions clarifies the robustness of the experience-description gap.
9. The advantage of I-SAW does not appear to be a result of the larger number of parameters. Some of the submitted reinforcement learning models have the same number of parameters as the best model. More

importantly, the competition method focuses on a prediction task, and for that reason addresses the risk of overfitting the data.

10. ACT-R (adaptive control of thought—rational) is general theory of cognition (see Anderson and Lebiere 1998).
11. Herrnstein et al. (1993, 150) write: "Melioration can be represented analytically as a type of partial maximization in which certain indirect effects are ignored or underweighted."
12. Nash equilibrium is defined as a prediction of the strategies of the different players from which no player has an incentive to deviate. That is, if a player believes that his or her opponent will follow a particular Nash prediction, he or she cannot benefit by deviating from this prediction. An equilibrium is weak if a deviation does not change the deviator's payoff.
13. Dal Bó (2005) ran repeated PD games with and without a fixed termination period. He found that in games without a fixed termination period, akin to infinitely repeated games, the "shadow of the future" significantly reduces opportunistic behavior.
14. It is important to stress that the summary of these results is based on using I-SAW as a benchmark and incorporating cognitive strategies to explain the observed deviations from the predictions of this benchmark. Different research methodologies may lead to other insights. Econometric investigation (e.g., Camerer and Ho 1999) can be useful, but insights tend to be sensitive to the assumption that the underlying model is well calibrated (see Feltovich 2000; Salmon 2001; Wilcox 2006; Erev and Haruvy 2005). Insights can also be derived from the long-term convergence properties of simple models (see Milgrom and Roberts 1990; Kalai and Lehrer 1993; Kandori, Mailath, and Rob 1993; Fudenberg and Levine 1998; Hart and Mas-Colell 2001), but insights are limited to situations with very long horizons and stationary payoffs.
15. The seating map can be used as evidence of cheating in the case of a disciplinary action to demonstrate that the students who have similar exam answers were also sitting next to one another.
16. By some accounts, the demise of Dr. Semmelweis was a function of his research (or correction) decisions. It seems that the influential heads of the departments who were responsible for the high and avoidable death rates were unhappy with his results.
17. We use the term *inventions* to refer to both naturally evolving institutions and to the outcomes of explicit mechanism design. The most important inventions, including the wheel, are the product of a process that includes natural evolution (e.g., people that rolled logs were more likely to survive) and some explicit design (e.g., the use of rubber to produce more effective wheels).

REFERENCES

Abbink, K., G. E. Bolton, A. Sadrieh, and F. F. Tang. 2001. Adaptive Learning versus Punishment in Ultimatum Bargaining. *Games and Economic Behavior* 37: 1–25.

Allais, M. 1953. Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américaine. *Econometrica* 21: 503–46.

Alm, James, and Michael McKee. 2004. Tax Compliance as a Coordination Game. *Journal of Economic Behavior & Organization* 54(3): 297–312.

Alos-Ferrer, Carlos, and Karl Schlag. 2009. Imitation and Learning. In P. Anand, P. Pattanaik, and C. Puppe, eds., *The Handbook of Rational and Social Choice*. Oxford, UK: Oxford University Press.

Anderson, J. R., and C. Lebiere. 1998. *The Atomic Components of Thought*. Mahwah, NJ: Erlbaum.

Anderson, Lisa R., and Charles A. Holt. 1997. Information Cascades in the Laboratory. *American Economic Review* 87(5): 847–62.

Andreoni, James, and J. H. Miller. 1993. Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence. *Economic Journal* 103(418): 570–85.

Apesteguia, Jose. 2006. Does Information Matter in the Commons? Experimental Evidence. *Journal of Economic Behavior and Organization* 60: 55–69.

Arifovic, J., R, D, McKelvey, and S. Pevnitskaya. 2006. An Initial Implementation of the Turing Tournament to Learning in Two Person Games. *Games and Economic Behavior* 57: 93–122.

Arifovic, Jasmina, and John Ledyard. 2004. Scaling Up Learning Models in Public Good Games. *Journal of Public Economic Theory* 6(2): 203–38.

Armantier, Olivier. 2004. Does Observation Influence Learning? *Games and Economic Behavior* 46: 221–39.

Ball, S., M. Bazerman, and J. Carroll. 1991. An Evaluation of Learning in the Bilateral Winner's Curse. *Organizational Behavior and Human Decision Processes* 48(1): 1–22.

Bandura, A. 1965. Influence of Models' Reinforcement Contingencies on the Acquisition of Imitative Response. *Journal of Personality and Social Psychology* 1: 589–95.

Barkan, R., D. Zohar, and I. Erev. 1998. Accidents and Decision Making under Uncertainty: A Comparison of Four Models. *Organizational Behavior and Human Decision Processes* 74: 118–44.

Baron, R., and M. Kalsher. 2006. *Psychology*. Boston, MA: Allyn and Bacon.

Barron, G., and I. Erev. 2003. Small Feedback-Based Decisions and Their Limited Correspondence to Description-Based Decisions. *Journal of Behavioral Decision Making* 16(3): 215–33.

Barron, G., S. Leider, and J. Stack. 2008. The Effect of Experience on a Warnings' Impact: Sex, Drugs and Rock-n-Roll. *Organizational Behavior and Human Decision Processes* 106: 125–42.

Barron, G., and E. Yechiam. 2009. The Coexistence of Overestimation and Underweighting of Rare Events and the Contingent Recency Effect. *Judgment and Decision Making* 4: 447–60.

Battalio, R. C., J. H. Kagel, and K. Jiranyakul. 1990. Testing between Alternative Models of Choice under Uncertainty: Some Initial Results. *Journal of Risk and Uncertainty* 3(1): 25–50.

Battalio, R., L. Samuelson, and J. Van Huyck. 2001. Optimization Incentives and Coordination Failure in Laboratory Stag Hunt Games. *Econometrica* 69: 749–64.

Bazerman, M. H., and M. A. Neale. 1992. The Mythical Fixed-Pie. *Executive Excellence* 9: 14–15.

Bechara, A., A. R. Damasio, H. Damasio, and S. W. Anderson. 1994. Insensitivity to Future Consequences Following Damage to Human Prefrontal Cortex. *Cognition* 50(1–3): 7–15.

Becker, G. 1968. Crime and Punishment: An Economic Approach. *The Journal of Political Economy* 76: 169–217.

Benartzi, S., and R. Thaler. 1995. Myopic Loss-Aversion and the Equity Premium Puzzle. *Quarterly Journal of Economics* 110(1): 73–92.

Ben Zion, U., I. Erev, E. Haruvy and T. Shavit. 2010. Adaptive Behavior Leads to Under-Diversification, *Journal of Economic Psychology* 31: 985–95.

Bereby-Meyer, Y., and I. Erev. 1998. On Learning to Become a Successful Loser: A Comparison of Alternative Abstractions of Learning Processes in the Loss Domain. *Journal of Mathematical Psychology* 42(2): 266–86.

Bereby-Meyer, Y., and B. Grosskopf. 2008. Overcoming the Winner's Curse: An Adaptive Learning Perspective. *Journal of Behavioral Decision Making* 21(1): 15–27.

Bereby-Meyer, Y., and A. Roth. 2006. The Speed of Learning in Noisy Games: Partial Reinforcement and the Sustainability of Cooperation. *American Economic Review* 96(4): 1029–42.

Berman, E. 2001. Sect, Subsidy and Sacrifice: An Economist's View of Ultra-Orthodox Jews. (In Hebrew). *The Economic Quarterly*.

Biele, G., I. Erev, and E. Ert. 2009. Learning, Risk Attitude and Hot Stoves in Restless Bandit Problems. *Journal of Mathematical Psychology* 53(3): 155–67.

Birnbaum, M. H., and T. Martin. 2003. Generalization across People, Procedures, and Predictions: Violations of Stochastic Dominance and Coalescing. In S. L. Schneider and J. Shanteau, eds., *Emerging Perspectives on Decision Research.* New York: Cambridge University Press, 84–107.

Blavatskyy, P. 2006. Axiomatization of a Preference for Most Probable Winner. *Theory and Design* 60: 17–33.

Blume, M. E., and I. Friend. 1975. The Asset Structure of Individual Portfolios and Some Implications for Utility Functions. *Journal of Finance* 30: 585–603.

Bolton, G., and A. Ockenfels. 2000. ERC: A Theory of Equity, Reciprocity and Competition. *American Economic Review* 90: 166–93.

Bornstein, G., D. V. Budescu, and S. Zamir. 1997. Cooperation in Intergroup, N-Person, and Two-Person, Games of Chicken. *Journal of Conflict Resolutions* 41: 384–406.

Brenner, T., and N. Vriend. 2006. On the Behavior of Proposers in the Ultimatum Game. *Journal of Economic Behavior and Organization* 61: 617–31.

Brewer, M. B., and D. T. Campbell. 1976. *Ethnocentrism and Intergroup Attitudes: East African Evidence*. New York: Sage.

Brown, G. W. 1951. Iterative Solution of Games by Fictitious Play. In T. Jallings and C. Koopmans, eds., *Activity Analysis of Production and Allocation*. New York: Wiley.

Brown, J. N., and R. Rosenthal. 1990. Testing the Minimax Hypothesis: A Re-Examination of O'Neill's Game Experiment. *Econometrica* 58(5): 1065–81.

Busemeyer, J. R., and I. J. Myung. 1988. A New Method for Investigating Prototype Learning. *Journal of Experimental Psychology: Learning, Memory, Cognition* 14: 3–11.

———. 1992. An Adaptive Approach to Human Decision Making: Learning Theory, Decision Theory, and Human Performance. *Journal of Experimental Psychology: General* 121: 177–94.

Busemeyer, J. R., and J. C. Stout, J. 2002. A Contribution of Cognitive Decision Models to Clinical Assessment: Decomposing Performance on the Bechara Gambling Task. *Psychological Assessment* 14: 253–62.

Busemeyer, J. R., and J. T. Townsend. 1993. Decision Field Theory: A Dynamic-Cognitive Approach to Decision Making in an Uncertain Environment. *Psychological Review* 100: 432–59.

Bush, R. R., and F. Mosteller. 1955. *Stochastic Models for Learning*. New York: Wiley.

Camerer, C. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.

Camerer, C., and T.-H. Ho. 1998. EWA Learning in Coordination Games: Probability Rules, Heterogeneity, and Time Variation. *Journal of Mathematical Psychology* 42: 305–26.

———. 1999. Experience-Weighted Attraction Learning in Normal Form Games. *Econometrica* 67(4): 827–74.

Camerer, C., T.-H. Ho, and J.-K. Chong 2002. Sophisticated Learning and Strategic Teaching. *Journal of Economic Theory* 104: 137–88.

Caplin, A., and M. Dean. 2007. The Neuroeconomic Theory of Learning. *American Economic Review Papers and Proceedings* 97(2): 148–52.

Charness, G., and B. Grosskopf. 2002. Cheap Talk, Information and Coordination-Experimental Evidence. Mimeo.

Charness, G., and M. Rabin. 2002. Understanding Social Preferences with Simple Tests. *Quarterly Journal of Economics* 117: 817–69.

Chen, W., S. Liu, C.-H. Chen, and Y.-S. Lee. 2011. Bounded Memory, Inertia, Sampling and Weighting Model for Market Entry Games. *Games 2*(1): 187–99.

Chen, Y., and F.-F. Tang. 1998. Learning and Incentive Compatible Mechanisms for Public Goods Provision: An Experimental Study. *Journal of Political Economy* 106: 633–62.

Cheung, Y.-W., and D. Friedman. 1997. Individual Learning in Normal Form Games: Some Laboratory Results. *Games and Economic Behavior* 19: 46–76.

———. 1998. Comparison of Learning and Replicator Dynamics Using Experimental Data. *Journal of Economic Behavior and Organization* 35: 263–80.

Clements, M. T. 2005. Inefficient Standard Adoption: Inertia and Momentum Revisited. *Economic Inquiry* 43(3): 507–18.

Coleman, A. A., A. M. Colman, and R. M. Thomas. 1990. Cooperation without Awareness: A Multiperson Generalization of the Minimal Social Situation. *Behavioral Science* 35: 15–121.

Colman, A. M. 2005. Evolution of Cooperation without Awareness in Minimal Social Situations. In N. Gold, ed., *Teamwork: Multi-Disciplinary Perspectives.* Basingstoke: Palgrave Macmillan, 216–35.

Colman, A. M., B. D. Pulford, D. Omtzigt, and A. al-Nowaihi. 2010. Learning to Cooperate without Awareness in Multiplayer Minimal Social Situations. *Cognitive Psychology* 61: 201–227.

Cooper, D., and J. Kagel. 2003. Lessons Learned: Generalizing Across Games. American *Economic Review* 93(2): 202–7.

———. 2008. Learning and Transfer in Signaling Games. *Economic Theory* 34(3): 415–40.

Crawford, G., and M. Shum. 2005. Uncertainty and Learning in Pharmaceutical Demand. *Econometrica* 37: 1137–74.

Cuchon, G., and C. Camerer. 1996. Loss Avoidance and Forward Induction in Coordination Games. *Quarterly Journal of Economics* 112: 165–94.

Dal Bó, P. 2005. Cooperation under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games. *American Economic Review* 95(5): 1591–1604.

Dal Bó, P., and G. Fréchette. 2011. The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence. *American Economic Review* 101: 411–29.

Daniely, H. 2000. Social Dilemmas and the Design of Simulators, MSc thesis, Technion, (in Hebrew).

Dayan, P., and Y. Niv. 2008. Reinforcement Learning: The Good, the Bad and the Ugly. *Current Opinion in Neurobiology* 18: 1–12.

Degroot, M. H. 1970. *Optimal Statistical Decisions*. New York: McGraw-Hill.

Delepoulle, S., P. Preux, and J.-C. Darcheville. 2000. Evolution of cooperation within a behavior-based perspective: Confronting nature and animats. Artificial Evolution Lecture Notes in Computer Science, 18291, 204–216.

Delepoulle, S., Preux, P., and Darcheville, J.-C. 2001. Selection of Behavior in Social Situations: Application to the Development of Coordinated Movements. *Applications of Evolutionary Computing, Proceedings* 2037: 384–93.

Dember, W. N., and H. Fowler. 1958. Spontaneous Alternation Behavior. *Psychological Bulletin* 55: 412–28.

Denrell, J. 2005. Why Most People Disapprove of Me: Experience Sampling in Impression Formation. *Psychological Review* 112(4): 951–78.

———. 2007. Adaptive Learning and Risk Taking. *Psychological Review* 114(1): 177–87.

Denrell, J., and G. Le Mens. 2007. Interdependent Sampling and Social Influence. *Psychological Review* 114(2): 398–422.

Denrell, J., and J. G. March. 2001. Adaptation as Information Restriction: The Hot Stove Effect. *Organization Science* 12(5): 523–38.

Diederich, A., and J. R. Busemeyer. 1999. Conflict and the Stochastic Dominance Principle of Decision Making. *Psychological Science* 10: 353–59.

Dufwenberg, M., T. Lindqvist, and E. Moore. 2005. Bubbles & Experience: An Experiment. *American Economic Review* 95: 1731–37.

Eck, J. E., and E. Maguire. 2000. Have Changes in Policing Reduced Violent Crime? An Assessment of the Evidence. In Alfred Blumstein and Joel Wallman, eds., *The Crime Drop in America*. New York: Cambridge University Press.

Edwards, W. 1961. Probability Learning in 1000 Trials. *Journal of Experimental Psychology* 62: 385–94.

Ehrblatt, W., K. Hyndman, E. Ozbay, and A. Schotter. 2006. Convergence: An Experimental Study of Teaching and Learning in Repeated Games. Working paper, New York University.

Einhorn, H. J., and R. M. Hogarth. 1978. Confidence in Judgment: Persistence of the Illusion of Validity. *Psychological Review* 85: 395–416

Engelbrecht-Wiggans, R., and E. Katok. 2007. Regret in Auctions: Theory and Evidence. *Economic Theory* 33: 81–101.

———. 2008. Regret and Feedback Information in First-Price Sealed-Bid Auctions. *Management Science* 54(4): 808–19.

Erdem, T., and M. Keane. 1996. Decision-Making under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets. *Marketing Science* 15(1): 1–20.

Erdem, T., M. Keane, T. Oncu, and J. Strebel. 2005. Learning about Computers: An Analysis of Information Search and Technology. *Quantitative Marketing and Economics* 3: 207–47.

Erev, I. 1998. Signal Detection by Human Observers: A Cutoff Reinforcement Learning Model of Categorization Decisions under Uncertainty. *Psychological Review* 105: 280–98.

———. 2007. On the Weighting of Rare Events and the Economics of Small Decisions. In S. H. Oda, ed., *Advances in Experimental Economics, Lecture Notes in Economics and Mathematical Systems*, Vol. 590. Dordrecht, The Netherlands: Springer.

Erev, I., and G. Barron. 2002. On Adaptation, Maximization, and Reinforcement Learning Among Cognitive Strategies. Working paper.

———. 2005. On Adaptation, Maximization, and Reinforcement Learning among Cognitive Strategies. *Psychological Review* 112: 912–31.

Erev, I., Y. Bereby-Meyer, and A. E. Roth. 1999. The Effect of Adding a Constant to All Payoffs: Experimental Iinvestigation, and Implications for Reinforcement Learning Models. *Journal of Economic Behavior and Organization* 39(1): 111–28.

Erev, I. E. Ert, A. E. Roth, E., Haruvy, S. Herzog, R. Hau, R. Hertwig, T. Stewart, R. West, and C. Lebiere. 2010a. A Choice Prediction Competition, for Choices from Experience and from Description. *Journal of Behavioral Decision Making* 23: 15–47.

———. 2010b. A Choice Prediction Competition for Market Entry Games: An Introduction. *Games* 1(2): 117–36.

Erev, I., E. Ert, and E. Yechiam. 2008. Loss Aversion, Diminishing Sensitivity, and the Effect of Experience on Repeated Decisions. *Journal of Behavioral Decision Making* 21(5): 575–97.

Erev, I., S. Gilat-Yihyie, D. Marchiori, and D. Sonsino. 2015. On Loss Aversion, Level-1 Reasoning, and Betting. *International Journal of Game Theory* 44(1): 113–33.

Erev, I., and D. Gopher. 1998. A Cognitive Game Theoretic Analysis of Attention Strategies, Ability and Incentives. *Attention and Performance* 17: 343–71.

Erev, I. and B. Greiner. 2015. The 1-800 Critique, Counter-Examples, and the Future of Behavioral Economics. In G. Fréchette and A. Schotter, eds., *The Methods of Modern Experimental Economics*. New York: Oxford University Press, 151–65.

Erev, I., and E. Haruvy. 2005. Generality and the Role of Descriptive Learning Models. *Journal of Mathematical Psychology* 49(5): 357–71.

Erev, I., P. Ingram, O. Raz, and D. Shany. 2010. Continuous Punishment and the Potential of Gentle Rule Enforcement. *Behavioural Processes* 84: 366–71.

Erev, I., and R. Livne-Tarandach. 2005. Experiment-Based Exams and the Difference between the Behavioral and the Natural Sciences. In R. Zwick and A. Rapoport, eds., *Experimental Business Research*, Vol 3. Dordrecht, The Netherlands: Springer.

Erev, I., and A. Rapoport. 1998. Coordination, "Magic", and Reinforcement Learning in a Market Entry Game. *Games and Economic Behavior* 23: 146–75.

Erev, I., and D. Rodansky. 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 and Commerce (in Hebrew).

Erev, I., D. Rodensky, M. Levi, M. Hershler, H. Adami, and Y. Donchin. 2010. The Value of Gentle Enforcement on Safe Medical Procedures. *Quality and Safety in Health Care* 19(5): 1–3.

Erev, I., and A. E. Roth. 1998. Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique Mixed Strategy Equilibria. *American Economic Review* 88: 848–81.

———. 1999. On the Role of Reinforcement Learning in Experimental Games: The Cognitive Game-Theoretic Approach. In D. Budescu, I. Erev, and R. Zwick, eds., *Games and Human Behavior: Essays in Honor of Amnon Rapoport. LEA*, 53–78.

———. 2001. Simple Reinforcement Learning Models and Reciprocation in the Prisoner's Dilemma Game. In G. Gigerenzer and R. Selten, eds., *Bounded Rationality: The Adaptive Toolbox*. Cambridge, MA: MIT Press, 215–31.

———. 2007. Multi-Agent Learning and the Descriptive Value of Simple Models. *Artificial Intelligence* 171(7): 423–28.

Erev, I., A. Roth, R. Slonim, and G. Barron. 2002. Predictive Value and the Usefulness of Game Theoretic Models. *International Journal of Forecasting* 18(3): 359–68.

———. 2007. Learning and Equilibrium as Useful Approximations: Accuracy of Prediction on Randomly Selected Constant Sum Games. *Economic Theory* 33: 29–51.

Erev, I., T. S. Wallsten, and D. V. Budescu. 1994. Simultaneous Over- and Underconfidence: The Role of Error in Judgment Processes. *Psychological Review* 101: 519–27.

Ert, E., and Y. Bereby-Meyer. Forthcoming. On Guessing, Choice and Probability Learning.

Ert E., and Erev, I. 2007. Replicated Alternatives and the Role of Confusion, Chasing, and Regret in Decisions from Experience. *Journal of Behavioral Decision Making* 20(3): 305–22.

———. 2008. The Rejection of Attractive Gambles, Loss Aversion, and the Lemon Avoidance Heuristic. *Journal of Economic Psychology* 29(5): 715–23.

———. 2016. It Won't Happen to Me": Review of Basic and Applied Studies of the Impact of Rare Events. In Vicki Bier, ed., *The Gower Handbook of Extreme Risk*. Farnham UK: Gower.

———. 2013. On the Descriptive Value of Loss Aversion in Decisions under Risk: Six Clarifications. *Judgment & Decision Making* 8(3).

Estes, W. K. 1950. Toward a Statistical Theory of Learning. *Psychological Review* 57: 94–107.

———. 1964. Probability Learning. In A. W. Melton, ed., *Categories of Human Learning*. New York: Academic Press, 89–128.

———. 1976. The Cognitive Side of Probability Learning. *Psychological Review* 83: 37–64.

Fabiani, M., J. Buckley, B. Gratton, M. G. H. Coles, E. Donchin, and R. Logie. 1989. The Training of Complex Task Performance. *Acta Psychologia* 71: 259–99.

Fehr, E., and K. M. Schmidt. 1999. A Theory of Fairness, Competition and Cooperation. *Quarterly Journal of Economics* 114(3): 817–68.

Feltovich, N. 2000. Reinforcement-Based vs. Beliefs-Based Learning in Experimental Asymmetric-Information Games. *Econometrica* 68(3): 605–41.

Feltovich, N., A. Iwasaki, and S. Oda. 2007. Payoff Levels, Loss Avoidance, and Equilibrium Selection in the Stag Hunt: An Experimental Study. Working paper.

Festinger, L., S. Schachter, and K. W. Back. 1950. *Social Pressures in Informal Groups: A Study of Human Factors in Housing*. Stanford, CA: Stanford University Press.

Fiedler K. 2000. Beware of Samples! A Cognitive-Ecological Sampling Approach to Judgment Biases. *Psychological Review* 107: 659–76.

Filiz-Ozbay, E., and E. Ozbay. 2007. Auctions with Anticipated Regret: Theory and Experiment. *American Economic Review* 97(4): 1407–18.

Friedman, D. 1989. The S-Shaped Value Function as a Constrained Optimum. *The American Economic Review* 79(5): 1243–48.

Friedman, D., and D. W. Massaro. 1998. Understanding Variability in Binary and Continuous Choice. *Psychonomic Bulletin and Review* 5: 370–89.

Friedman, M., and L. J. Savage. 1948. The Utility Analysis of Choices Involving Risk. *Journal of Political Economy* 56(4): 279–304.

Fudenberg, D., and D. K. Levine. 1998. *The Theory of Learning in Games*. Cambridge, MA: MIT Press.

Fujikawa, T. 2009. On the Relative Importance of the Hot Stove Effect and the Tendency to Rely on Small Samples. *Judgment and Decision Making* 4: 429–35.

Fujikawa, T., and S. H. Oda. 2007. Judgment in Small Decision-Making Problems. In S. H. Oda, ed., *Developments on Experimental Economics*. Berlin, Germany: Springer Verlag.

Gaissmaier, W., and L. J. Schooler. 2008. The Smart Potential behind Probability Matching. *Cognition* 109: 416–22. doi:10.1016/j.cognition.2008.09.007.

Gans, N., G. Knox, and R. Croson. 2007. Simple Models of Discrete Choice and Their Performance in Bandit Experiments. *Manufacturing and Service Operations Management* 9: 383–408.

Garcia, J., D. J. Kimeldorf, and R. A. Koelling. 1955. Conditioned Aversion to Saccharin Resulting from Exposure to Gamma Radiation. *Science, New Series* 122(3160): 157–58.

Garcia, J., and T. A. Koelling. 1966. A Relation of Cue to Consequence in Avoidance Learning. *Psychonomic Science* 4: 123–24.

Gavetti, G., and D. Levinthal. 2000. Looking Forward and Book Backward: Cognitive and Experiential Search. *Administrative Science Quarterly* 45: 113–37.

Gershman, S. J., D. M. Blei, and Y. Niv. 2010. Context, Learning and Extinction. *Psychological Review* 117(1): 197–209.

Gigerenzer, G., and R. Selten. 2001. *Bounded Rationality: The Adaptive Toolbox*. London: MIT Press.

Gittines, J. C. 1979. Bandit Processes and Dynamic Allocation Indices. *Journal of the Royal Statistical Society, Series B* 41: 148–77.

———. 1989. *Multi-Armed Bandit Allocation Indices*. Hoboken, NJ: Wiley.

Glines, C. V. 2005. Top Secret WWII Bat and Bird Bomber Program. *Aviation History* 15(5): 38–44.

Golub, A., B. Johnson, A. Taylor, and H. Liberty. 2002. The Validity of Arrestee's Self Reports. *Justice Quarterly* 19(3): 477–502.

Gonzalez, C., V. Dutt, and T. Lejarraga. 2011. A Loser Can Be a Winner: Comparison of Two Instance-Based Learning Models in a Market Entry Competition. *Games* 2(1): 136–62.

Gonzalez, C., J. F. Lerch, and C. Lebiere. 2003. Instance-Based Learning in Dynamic Decision Making. *Cognitive Science* 27: 591–635.

Gopher, D., M. Weil, and T. Bereket. 1994. Transfer of Skill from a Computer Game Trainer to Flight. *Human Factors* 33(3): 387–405.

Gopher, D., M. Weil, and D. Siegal. 1989. Practice under Changing Priorities: An Approach to the Training of Complex Skills. *Acta Psychologica* 71: 147–77.

Gourville, J. T. 2003. Why Consumers Don't Buy: The Psychology of New Product Adoption. Harvard Business School Case No. 504–056.

Graham, G. H., and J. Unruh. 1990. The Motivational Impact of Non-Financial Employee Appreciation Practices on Medical Technologists. *Health Care Supervisor* 8(3): 9–17.

Grant, D. A., H. W. Hake, and J. P. Hornseth. 1951. Acquisition and Extinction of a Verbal Conditioned Response with Differing Percentages of Reinforcement. *Journal of Experimental Psychology* 42: 1–5.

Green, L., E. B. Fisher, Jr., S. Perlow, and L. Sherman. 1981. Preference Reversal and Self Control: Choice as a Function of Reward Amount and Delay. *Behavior Analysis Letters* 2: 103–12.

Grinblatt, M., S. Titman, and R. Wermers. 1995. Momentum Investment Strategies Portfolio Performance, and Herding: A Study of Mutual Fund Behavior. *American Economic Review* 85: 1088–1105.

Grosskopf, B. 2003. Reinforcement and Directional Learning in the Ultimatum Game with Responder Competition. *Experimental Economics* 6(2): 141–58.

Grosskopf, B., I. Erev, and E. Yechiam. 2006. Foregone with the Wind: Indirect Payoff Information and Its Implications for Choice. *International Journal of Game Theory* 34(2): 285–302.

Guth, W., N. Marchand, and J.-L. Rulliere. Forthcoming. On the Reliability of Reciprocal Fairness—An Experimental Study. *Eastern Economic Review*.

Guth, W., K. Schmittberger, and B. Schwarze. 1982. An Experimental Analysis of Ultimatum Bargaining. *Journal of Economic Behavior and Organization* 3(4): 367–88.

Guth, W., and E. Van Damme. 1998. Information, Strategic Behavior and Fairness in Ultimatum Bargaining—An Experimental Study. *Journal of Mathematical Psychology* 42: 227–47.

Hanaki, N., R. Sethi, I. Erev, and A. Peterhansl. 2005. Learning Strategies. *Journal of Economic Behavior and Organization* 56(4): 523–42.

Harcourt, B. G. 2001. *Illusion of Order: The False Premise of Broken Windows Policing*. Cambridge, MD: Harvard University Press.

Harrison, Glenn W. 1989. Theory and Misbehavior of First-Price Auctions. *American Economic Review* 79: 749–62.

Hart, S., and A. Mas-Colell. 2001. A General Class of Adaptive Strategies. *Journal of Economic Theory* 98: 26–54.

Haruvy, E., and I. Erev. 2001. On the Application and Interpretation of Learning Models. In R. Rami Zwick and A. Rapoport, eds., *Experimental Business Research*. Boston: Kluwer Academic Publishers, 285–300.

Haruvy E., I. Erev, and D. Sonsino. 2001. The Medium Prizes Paradox: Evidence from a Simulated Casino. *Journal of Risk and Uncertainty* 22: 251–61.

Haruvy, E., Y. Lahav, and C. Noussair. 2007. Traders' Expectations in Asset Markets: Experimental Evidence. *American Economic Review* 97(5): 1901–20.

Hau, R., T. Pleskac, and R. Hertwig. 2008. Decisions from Experience without Sampling Error. *Journal of Behavioral Decision Making* 21: 493–518.

Healy, A. E., and M. Kubovy. 1981. Probability Matching and the Formation of Conservative Decision Rules in a Numerical Analog of Signal Detection. *Journal of Experimental Psychology: Human Learning and Memory* 7: 344–54.

Herrnstein, R. J. 1961. Relative and Absolute Strength of Responses as a Function of Frequency of Reinforcement. *Journal of the Experimental Analysis of Behavior* 4: 267–72.

———. 1988. A Behavioral Alternative to Utility Maximization. In S. Maital, ed., *Applied Behavioral Economics*. London: Wheatsheaf Books.

Herrnstein, R. J., G. F. Lowenstein, D. Prelec, and W. Vaughan Jr., 1993. Utility Maximization and Melioration: Internalities in Individual Choice. *Journal of Behavioral Decision Making* 6: 149–85.

Herrnstein, R. J., and J. Mazur. 1987. Making Up Our Minds. *The Sciences* 27: 40–47.

Herrnstein, R. J., and D. Prelec. 1991. Melioration: A Theory of Distributed Choice. *The Journal of Economic Perspectives* 5(3): 137–56.

Herrnstein, R. J., and W. Vaughan Jr. 1980. Melioration and Behavioral Allocation. In J. E. Stadden, ed., *Limits to Action*. New York: Academic Press.

Hertwig, R., G. Barron, E. U. Weber, and I. Erev. 2004. Decisions from Experience and the Effect of Rare Events in Risky Choice. *Psychological Science* 15: 534–39.

Hertwig, R., and I. Erev. 2009, The Description–Experience Gap in Risky Choice. *Trends in Cognitive Sciences* 13: 517–23.

Hertwig, R., and A. Ortmann. 2002. Economists' and Psychologists' Experimental Practices: How They Differ, Why They Differ and How They Could Converge. In I. Brocas, and J. D. Carillo, eds., *The Psychology of Economic Decisions*. New York: Oxford University Press, 253–72.

Hertwig, R., and T. J. Pleskac. 2010. Decisions from Experience: Why Small Samples? *Cognition* 115(2): 225–37.

Ho, T.-H., C. Camerer, and J.-K. Chong. 2007. Self-Tuning Experience Weighted Attraction Learning in Games. *Journal of Economic Theory* 133(1): 177–98.

Ho, T-H., Camerer, C., and Weigelt, K. 1998. Iterated Dominance and Iterated Best Response in Experimental P-Beauty Contests. *The American Economic Review* 88: 947–69.

Ho, T-H, X. Wang, and C. Camerer. 2008. Individual Differences in the EWA Learning with Partial Payoff Information. *The Economic Journal* 118: 37–59.

Ho, T.-H. and K. Weigelt. 1996. Task Complexity, Equilibrium Selection, and Learning: An Experimental Study. *Management Science* 42: 659–79.

Hochman, G., and I. Erev. 2013 The Partial Reinforcement Extinction Effect and the Contingent Sampling Hypothesis. *Psychonomics Bulletin and Review* 20(6): 1336–42.

Hopkins, E. 2002. Two Competing Models of How People Learn in Games. *Econometrica* 70(6): 2141–66.

Huck, S., H.-T. Normann, and J. Oechssler. 1999. Learning in Cournot Oligopoly—An Experiment. *The Economic Journal* 109: 80–95.

Huck, S., J. Oechssler, and H. T. Normann. 2003. Zero Fnowledge Cooperation in Dilemma Games. *Journal of Theoretical Biology* 220: 47–54.

Humphreys, L. G. 1939a. The Effect of Random Alternation of Reinforcement on the Acquisition and Extinction of Conditioned Eyelid Reactions. *Journal of Experimental Psychology* 25: 141–58.

———. 1939b. Acquisition and Extinction of Verbal Expectations in a Situation Analogous to Conditioning. *Journal of Experimental Psychology* 25: 294–301.

Isaac, M., and C. Plott. 1981. Price Controls and the Behavior of Auction Markets: An Experimental Examination. *American Economic Review* 71(3): 448–59.

Isaac, R. M., and J. M. Walker. 1988. Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism. *Quarterly Journal of Economics* 103: 179–99.

Ivanova-Stenzel, R., and T. Salmon. 2004. Bidder Preferences among Auction Institutions. *Economic Inquiry* 42: 223–36.

Jessup, R.K., A. J. Bishara, and J. R. Busemeyer. 2008. Feedback Produces Divergence from Prospect Theory in Descriptive Choice. *Psychological Science* 19(10): 1015–22.

John, E. R., P. Chesler, F. Bartlett, and I. Victor. 1969. Observational Learning in Cats. *Science* 166: 901–3.

Juvina, I., M. Saleem, J. M. Martin, C. Gonzalez, and C. Lebiere. 2013. Reciprocal Trust Mediates Deep Transfer of Learning between Games of Strategic Interaction. *Organizational Behavior and Human Decision Processes* 120(2): 206–15.

Kagel, J. H., R. Battalio, and L. Green. 1995. *Economic Choice Theory: An Experimental Analysis of Animal Behavior*. New York: Cambridge University Press.

Kagel, J. H., R. M. Harstad, and D. Levin. 1987. Information Impact and Allocation Rules in Auctions with Affiliated Private Values: A Laboratory Study. *Econometrica* 55(6): 1275–04.

Kagel, J. H., and D. Levin. 1986. The Winner's Curse and Public Information in Common Value Auctions. *American Economics Review* 76: 894–920.

Kagel, J. H., and P. McGhee, forthcoming, Team versus Individual Play in Finitely Repeated Prisoner Dilemma Games. *AEJ: Microeconomics*.

Kahneman, D. 1988. Experimental Economics: A Psychological Perspective. In R. Tietz, W. Albers, and R. Selten, eds., *Bounded Rational Behavior in Experimental Games and Markets*. Berlin: Springer-Verlag.

Kahneman, D., and A. Tversky. 1979. Prospect Theory: An Analysis of Decisions under Risk. *Econometrica* 47: 313–27.

Kalai, E., and E. Lehrer. 1993. Rational Learning Leads to Nash Equilibrium. *Econometrica* 61(5): 1019–45.

Kandori, M., G. Mailath, and R. Rob. 1993. Learning, Mutation, and Long Run Equilibria in Games. *Econometrica* 61(1): 29–56.

Kareev, Y. 2000. Seven (Indeed, Plus or Minus Two) and the Detection of Correlations. *Psychological Review* 107(2): 397.

Karpoff, J. M. 1988. Costly Short Sales and the Correlation of Returns with Volume. *Journal of Financial Research* 11(3): 173–88.

Kelling, G., and W. H. Sousa Jr. 2001. Do Police Matter? An Analysis of the Impact of New York City's Police Reforms. Manhattan Institute Center for Civic Innovation Civic Report No. 22.

Kelling G., and J. Q. Wilson. 1982. Broken Windows: The Police and Neighborhood Safety. *Atlantic Monthly* 249: 29–38.

Kelly, M. 1995. All Their Eggs in One Basket: Portfolio Diversification of U.S. Households. *Journal of Economic Behavior and Organization* 27: 87–96.

Kliger, D., O. Levy, and D. Sonsino. 2003. On Absolute and Relative Performance and the Demand for Mutual Funds—Experimental Evidence. *Journal of Economic Behavior and Organization* 52: 341–63.

Lebiere, C., C. Gonzalez, and M. Martin. 2007. Instance-Based Decision Making Model of Repeated Binary Choice. In *Proceedings of the 8th International Conference on Cognitive Modeling*. Ann Arbor, MI.

Lebiere, C., D. Wallach, and R. L. West. 2000. A Memory-Based Account of the Prisoner's Dilemma and Other 2×2 Games. In *Proceedings of International Conference on Cognitive Modeling*. NL: Universal Press, 185–93.

Ledyard, J. O. 1995. Public Goods: A Survey of Experimental Research. In J. Kagel and A. Roth, eds., *The Handbook of Experimental Economics*, Vol. 1. Princeton, NJ: Princeton University Press.

Lee, W. 1971. *Decision Theory and Human Behavior*. New York: Wiley.

Lejarraga, T., and C. Gonzalez. 2011. Effects of Feedback and Complexity on Repeated Decisions from Description. *Organizational Behavior and Human Decision Processes* 116(2): 286–95.

Lemieux, T., and W. B. MacLeod. 2000. Supply Side Hysteresis: The Case of the Canadian Unemployment Insurance System. *Journal of Public Economics* 78(1–2): 139–70.

Levinthal, D. A., and J. G. March. 1993. The Myopia of Learning. *Strategic Management Journal, Winter Special Issue* 14: 95–112.

List, J. A., and T. L. Cherry. 2003. Examining the Role of Fairness in High Stakes Allocation Decisions. Working paper.

Lovio-George, C. 1992. What Motivates Best? *Sales & Marketing Management* (April).

Ludvig E. A., and M. L. Spetch. 2011. Of Black Swans and Tossed Coins: Is the Description-Experience Gap in Risky Choice Limited to Rare Events? *PLoS ONE* 6(6): e20262.

Luria, A., I. Erev, and E. Haruvy. 2016. The Reinforcing Value of Lottery Tickets, and the Synergistic Effect of Distinct Reinforcements. *Journal of Behavioral Decision Making*.

Luria, G., D. Zohar, and I. Erev. 2008. The Effect of Visibility on Effectiveness of Supervisory-Based Interventions. *Journal of Safety Research* 39: 273–80.

MacDonald, D., J. Kagel, and R. Battalio. 1991. Animals' Choices over Uncertain Outcomes: Further Experimental Results. *Economic Journal* 191: 1067–84.

Macy, M. W., and A. Flache. 2002. Learning Dynamics. In *Social Dilemmas, Proceedings of the National Academy of Sciences of the United States of America* 99(10), Supplement 3.

Maier, S. F., and M. E. P. Seligman. 1976. Learned Helplessness: Theory and Evidence. *Journal of Experimental Psychology: General* 105: 3–46.

Malcom, D., and B. Liebermann. 1965. The Behavior of Responsive Individuals Playing a Two-Person, Zero-Sum Game Requiring the Use of Mixed Strategies. *Psychonomic Science*: 373–74.

Mane, A. M., and E, Donchin. 1989. The Space Fortress Game. *Acta Psychologica* 71: 17–22.

Marchiori, D., S. Di Guida, and I. Erev 2015. Noisy Retrieval Models of Over- and Undersensitivity to Rare Events. *Decision* 2(2): 82.

Marchiori, D., and M. Warglien. 2008. Predicting Human Interactive Learning by Regret-Driven Neural Networks. *Science* 319: 1111–1113.

McFadden, D. L. 1974. Conditional Logit Analysis of Qualitative Choice Behavior. In Paul Zarembka, ed., *Frontiers in Econometrica*. New York: Academic Press, 105–42.

McKelvey, R. D., and T. R. Palfrey. 1995. Quantal Response Equilibrium for Normal Form Games. *Games and Economic Behavior* 10: 6–38.

———. 2001. Playing in the Dark: Information, Learning and Coordination in Repeated Games. Technical report, California Institute of Technology.

Merlo, A., and A. Schotter. 1992. Experimentation and Learning in Laboratory Experiments: Harrison's Criticism Revisited. *American Economic Review* 82(5): 1413–25.

———. 2003. Learning By Not Doing: An Experimental Study. Games and Economic *Behavior* 42: 116–36.

Meyer, R. J., and Y. Shi. 1995. Sequential Choice under Ambiguity: Intuitive Solutions to the Armed Bandit Problem. *Management Science* 41: 817–34.

Milgrom, P., and J. Roberts. 1990. Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities. *Econometrica* 58(6): 1255–77.

Milgrom, Paul. 2004. *Putting Auction Theory to Work*. New York: Cambridge University Press.

Miller, N. F., and J. Dollard. 1941. *Social Learning and Imitation*. New Haven: Yale University Press.

Mitropoulos, A. 2001. Learning under Minimal Information: An Experiment on Mutual Fate Control. *Journal of Economic Psychology* 22: 523–57.

———. 2003. An Experiment on the Value of Structural Information in a 2×2 Repeated Game. *Economics Letters* 78: 27–32.

Mookherjee, D., and B. Sopher. 1994. Learning Behavior in an Experimental Matching Pennies Game. *Games and Economic Behavior* 7: 62–91.

———. 1997. Learning and Decision Costs in Experimental Constant Sum Games. *Games and Economic Behavior* 19(1): 97–132.

Mowrer, O. H. 1947. On the Dual Nature of Learning—A Reinterpretation of “Conditioning” and “Problem Solving.” *Harvard Educational Review* 17: 102–48.
Myers, J., and E. Sadler. 1960. Effects of Range of Payoffs as a Variable in Risk Taking. *Journal of Experimental Psychology* 60: 306–09.
Myers, J., M. Suydam, and B. Gambino. 1965. Contingent Gains and Losses in Risk Taking Situations. *Journal of Mathematical Psychology* 2: 363–70.
Nash, J. 1950. Non-Cooperative Games. PhD thesis, Princeton University.
Nelson, B. 1994. *1001 Ways to Energize Employees*. New York: Workman Publishing Company.
———. 1999. The Use of Informal Rewards in Recognizing Performance. http://www.p-management.com/articles/9902.htm.
Neugebauer, T., and R. Selton. 2006. Individual Behavior of First-Price Auctions: The Importance of Information Feedback in Computerized Experimental Markets. *Games and Economic Behavior* 54(1): 183–204.
Nevin J. A. 1988. Behavioral Momentum and the Partial Reinforcement Effect. *Psychological Bulletin*. 103: 44–56.
Nevo, I., and I. Erev. 2012. On Surprise, Change, and the Effect of Recent Outcomes. *Frontiers in Cognitive Science*.
Nuland, S. B. 2003. *The Doctors’ Plague: Germs, Childbed Fever and the Strange Story of Ignac Semmelweis*. New York: W. W. Norton.
Nyarko, Y., and A. Schotter. 2002. An Experimental Study of Belief Learning Using Elicited Beliefs. *Econometrica* 70: 971–1006.
Ochs, Jack. 1995. An Experimental Study of Games with Unique Mixed Strategy Equilibria. *Games and Economic Behaviour* 202–17.
Offerman, T., J. Potters, and J. Sonnemans. 2002. Imitation and Belief Learning in an Oligopoly Experiment. *Review of Economic Studies* 69(4): 973–97.
O’Neill, B. 1987. Nonparametric Test of the Minimax Theory of Two-Person Zero-Sum Games. *Proceedings of the National Academy of Sciences* 84(7): 2106–9.
Osborne, M., and A. Rubinstein. 1998. Games with Procedurally Rational Players. *American Economic Review* 88: 834–47.
Osgood, C. E. 1949. The Similarity Paradox in Human Learning: A Resolution. *Psychological Review* 56: 132–43.
Overmier, J. B., and M.E.P. Seligman. 1967. Effects of Inescapable Shock upon Subsequent Escape and Avoidance Responding. *Journal of Comparative and Physiological Psychology* 63: 28–33.
Pavlov, I. P. 1927. *Conditioned Reflexes*. London: Oxford University Press.
Payne, J. W., J. R. Bettman, and E. L. Johnson. 1993. *The Adaptive Decision Maker*. Cambridge, UK: Cambridge University Press.
Petersen, C., S. F. Maier, and M.E.P. Seligman. 1995. *Learned Helplessness: A Theory for the Age of Personal Control*. New York: Oxford University Press.
Plonsky, O., K. Teodorescu, and I. Erev. 2015. Reliance on Small Samples, the Wavy Recency Effect, and Similarity-Based Learning. *Psychological Review* 122(4): 621–47.
Plott, C. R. 1982. Industrial Organization Theory and Experimental Economics. *Journal of Economic Literature* 20: 1485–1527.
Plott, C. R., and V. L. Smith. 1978. An Experimental Examination of Two Exchange Institutions. *Review of Economic Studies* 45: 133–53.
Radner, R. 1980. Collusive Behavior in Non-Cooperative Epsilon Equilibria of Oligopolies with Long but Finite Lives. *Journal of Economic Theory* 22: 121–57.
Rakow, T., and B. R. Newell. 2010. Degrees of Uncertainty: An Overview and Framework for Future Research on Experience-Based Choice. *Journal of Behavioral Decision Making* 23: 1–14.
Rapoport, A. 1988. Provision of Step-Level Public Goods: Effects of Inequality in Resources. *Journal of Personality and Social Psychology* 54: 432–40.
———. 1995. Effects of Information on Assessment of Probabilities. *Theory and Decision* 41(21).

Rapoport, A., and R. Boebel. 1992. Mixed Strategies in Strictly Competitive Games: A Further Test of the Minimax Hypothesis. *Games and Economic Behavior* 4: 261–83.

Rapoport, A., and D. V. Budescu. 1997. Randomization in Individual Choice Behavior. *Psychological Review* 104: 603–17.

Rapoport, A., T. E. Daniel, and D. A. Seale. 1998. Reinforcement-Based Adaptive Learning in Asymmetric Two-Person Bargaining with Incomplete Information. *Journal of Experimental Economics* 1(3): 221–53.

Rapoport, A., I. Erev, E. V. Abraham, and D. E. Olson. 1997. Randomization and Adaptive Learning in a Simplified Poker Game. *Organizational Behavior and Human Decision Processes* 69: 31–49.

Rapoport, A., and A. Mowshowitz. 1966. Experimental Studies of Stochastic Models for the Prisoner's Dilemma. *Behavioral Science* 11: 444–58.

Rapoport, A., D. A. Seale, and L. Ordonez. 2002. Tacit Coordination in Choice Between Certain Outcomes in Endogenously Determined Lotteries. *The Journal of Risk and Uncertainty* 25(1): 21–45.

Rapoport, A., D. A. Seale, and E. Winter. 2002. Coordination and Learning Behavior in Large Groups with Asymmetric Players. *Games and Economic Behavior* 39: 11–136.

Rapoport, A., and J. Sundali. 1996. Ultimatums in Two Person Bargaining with One Sided Uncertainty: Offer Games. *International Journal of Game Theory* 25: 475–94.

Rapoport, An. 1967. Exploiter, Leader, Hero, and Martyr: The Four Archetypes of the 2 × 2 Game. *Behavioral Science* 12: 81–84.

Rapoport, An., and A. M. Chammah. 1965. *Prisoner's Dilemma*. Ann Arbor, MI: University of Michigan Press.

Rapoport, An., M. J. Guyer, and D. G. Gordon. 1976. *The* 2 × 2 *Game*. Ann Arbor, MI: University of Michigan Press.

Redish, A., S. Jensen, S. Johnson, and A. Kurth-Nelson. 2007. Reconciling Reinforcement Learning Models with Behavioral Extinction and Renewal: Implications for Addiction, Relapse, and Problem Gambling. *Psychological Review* 114: 784–805.

Rescorla, R. A., and R. L. Solomon. 1967. Two-Process Learning Theory: Relationships between Pavlovian Conditioning and Instrumental Learning. *Psychological Review* 74: 151–82.

Richman, C. L., W. N. Dember, and P. Kim. 1986. Spontaneous Alternation Behavior in Animals: A Review. *Current Psychology* 5(4): 358–91.

Rieskamp, J., and P. E. Otto. 2006. SSL: A Theory of How People Learn to Select Strategies. *Journal of Experimental Psychology: General* 135(2): 207–36.

Rieskamp, Jerome, R. Busemeyer, and Tei Laine. 2003. How Do People Learn to Allocate Resources? Comparing Two Learning Theories. *Journal of Experimental Psychology. Learning, Memory, and Cognition* 29(6): 1066–81.

Robbins, S. P. 2001. *Organizational Behavior*. Upper Saddle River, NJ: Prentice Hall.

Robinson, J. 1951. An Iterative Method of Solving a Game. *Annals of Mathematics* 54, 296–301.

Roth, A. E., and I. Erev. 1995. Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term. *Games and Economic Behavior* 8: 164–212.

Roth, A. E., V. Prasnikar, M. Okuwo-Fujiwara, and S. Zamir. 1991. Bargaining and Market Behavior in Jerusalem, Ljubiljana, Pittsburgh and Tokyo: An Experimental Study. *American Economic Review* 81: 1068–95.

Rottenstreich, Y. 1995. Framing in Games. Unpublished doctoral dissertation, Stanford University, Stanford, California.

Salmon, T. 2001. An Evaluation of Econometric Models of Adaptive Learning. *Econometrica* 69(6): 1597–1628.

Samuelson, P. 1963. Risk and Uncertainty: A Fallacy of Large Numbers. *Scientia* 9: 108–13.

Samuelson, W., and M. Bazerman. 1985. The Winner's Curse in Bilateral Negotiations. In V. L. Smith, ed., *Research in Experimental Economics*, Vol. 3. Greenwich, CT: JAI Press, 105–37.

Samuelson, W., and T. Zeckhauser. 1988. Status Quo Bias in Decision Making. *Journal of Risk and Uncertainty* 1: 7–59.

Sarin, R. and F. Vahid. 1999. Payoff Assessments without Probabilities: A Simple Dynamic Model of Choice. *Games and Economic Behavior* 28: 294–309.

———. 2001. Predicting How People Play Games: A Simple Dynamic Model of Choice. *Games and Economic Behavior* 34(1): 104–22.

Sutton, R. S., and A. G. Barto. 1998. *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Schmalensee, R. 1975. Alternative Models of Bandit Selection. *Journal of Economic Theory* 10: 333–42.

Schultz, W. 1998. Predictive Reward Signal of Dopamine Neurons. *Journal of Neurophysiology* 80(1): 1–27.

Schurr, A., D. Rodensky, and I. Erev. 2014. The Effect of Unpleasant Experiences on Evaluation and Behavior. *Journal of Economic Behavior & Organization* 106: 1–9.

Segal, M. W. 1974. Alphabet and Attraction: An Unobtrusive Measure of the Effect of Propinquity in a Field Setting. *Journal of Personality and Social Psychology* 30: 654–57.

Selten, R., and R. Stoecker. 1986. End Behaviour in Sequences of Finite Prisoner's Dilemma Supergames: A Learning Theory Approach. *Journal of Economic Behavior & Organization* 7: 47–70.

Shachat, J. M. 2002. Mixed Strategy Play and the Minimax Hypothesis. *Journal of Economic Theory* 104(1): 189–226.

Shafir, S. 2000. Risk-Sensitive Foraging: The Effect of Relative Variability. *Oikos* 88: 663–69.

Shafir, S.,T. Reich, E. Tsur, I. Erev, and A. Lotem. 2008. Perceptual Accuracy and Conflicting Effects of Certainty on Risk-Taking Behaviour. *Nature* 453: 917–20.

Shanks, D. R., R. J. Tunney, and J. D. McCarthy.2002. A Re-Examination of Probability Matching and Rational Choice. *A Journal of Behavioral Decision Making* 15(3): 232–50.

Sherman, R. 1971. Empirical Oligopoly. *Kyklo* 24(1): 30–49.

Shteingart, H., T. Neiman, and Y. Loewenstein. 2013. The Role of First Impression in Operant Learning. *Journal of Experimental Psychology: General* 142(2): 476.

Sidowski, J. B., L. B. Wyckoff, and L. Tabory. 1956. The Influence of Reinforcement and Punishment in a Minimal Social Situation. *Journal of Abnormal and Social Psychology* 52: 115–19.

Siegel, S., and D. A. Goldstein. 1959. Decision-Making Behavior in a Two-Choice Uncertain Outcome Situation. *Journal of Experimental Psychology* 57: 37–42.

Silverman, E. B. 1999. *New York Police Department Battles Crime*. Boston: Northeastern University Press.

Simon, H. 1957. Models of Man: Social and Rational. New York: Wiley.

Simonsohn, U., N. Karlsson, G. Loewenstein, and D. Ariely. Forthcoming. The Tree of Experience in the Forest of Information: Overweighing Experienced Relative to Observed Information. *Games and Economic Behavior*.

Skinner, B. F. 1938. *The Behavior of Organisms*. New York: Appleton-Century-Crofts.

Skogan, W. G. 1990. *Disorder and Decline: Crime and the Spiral of Decay in American Neighborhoods*. Berkeley, CA: University of California Press.

———. 1953. *Science and Human Behavior*. New York: Free Press.

Slonim R. 2007. Empirical Regularities in Behavior across a Population of Games. Working paper.

Slonim, R., and A. E. Roth. 1998. Learning in High Stakes Ultimatum Games: An Experiment in the Slovak Republic. *Econometrica* 66(3): 569–96.

Smalley, W., D. Shatin, D. K. Wysowski, J. Gurwitz, S. E. Andrade, M. Goodman, K. A. Chen, R. Platt, S. D. Schech, and W. A. Ray. 2000. Contraindicated Use of Cisapride: Impact of Food and Drug Administration Regulatory Action. *Journal of the American Medical Association* 284(23): 3036–39.

Smith, T. G., and A. Tasnádi. 2007. A Theory of Natural Addiction. *Games and Economic Behavior* 59: 316–44.

Smith, V. L. 1962. An Experimental Study of Competitive Market Behavior. *Journal of Political Economics* 70: 111–37.

Smith, V. L., G. L. Suchanek, and A. W. Williams. 1988. Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets. *Econmetrica* 56(5): 1119–51.

Sonsino, D. 1997. Learning to Learn, Pattern Recognition and Nash Equilibrium. *Games and Economic Behavior* 18: 286–331.

Sonsino, D., I. Erev, and S. Gilat, 2002. On the Likelihood of Repeated Zero-Sum Betting by Adaptive (Human) Agents. Mimeo.

Stahl, D. 1996. Boundedly Rational Rule Learning in a Guessing Game, *Games and Economic Behavior* 16: 303–30.

———. 1999. Evidence Based Rules and Learning in Symmetric Normal-Form Games, *International Journal of Game Theory* 28: 111–30.

———. 2000. Rule Learning in Symmetric Normal-Form Games: Theory and Evidence. *Games and Economic Behavior* 32: 105–38.

Stajkovic, A. and P. Luthans. 1997. A Meta-Analysis of the Effects of Organizational Behavior Modification on Task Performance, 1975–95. *Academy of Management Journal* 40: 1122–49.

Sterman, J. 1989. Misperceptions of Feedback in Dynamic Decision Making. *Organizational Behavior and Human Decision Processes* 43: 301–35.

Straus, M. A. 1991. Discipline and Deviance: Physical Punishment of Children and Violence and Other Crime in Adulthood. *Social Problems* 38(2): 133–54.

Sundali, J. A., A. Rapoport, and D. A. Seale. 1995. Coordination in Market Entry Games with Symmetric Players. *Organizational Behavior and Human Decision Processes* 64: 203–18.

Suppes, P., and Richard C. Atkinson. 1960. *Markov Learning Models for Multiperson Interactions*. Stanford, CA: Stanford University Press.

Sutherland, N. S., and N. J. Mackintosh. 1971. *Mechanisms of Animal Discrimination Learning*. New York: Academic Press.

Taleb, N. N. 2007. *The Black Swan: The Impact of the Highly Improbable*. New York: Random House.

Teodorescu, K., and I. Erev, 2014a. On the Decision to Explore New Alternatives: The Coexistence of Under-and Over-exploration. *Journal of Behavioral Decision Making* 27(2): 109–23.

———. 2014b. Learned Helplessness and Learned Prevalence: Exploring the Causal Relations of Perceived Controllability, Reward Prevalence and Exploration. *Psychological Science* 25(10): 1861–69.

Terkel, J. 1996. Cultural Transmission of Feeding Behavior in Black Rats (Rattus rattus). In C. Heynes and B. Galef Jr., eds., *Social Learning in Animals and the Roots of Culture*. New York: Academic Press, 17–47.

Thaler, R. H., A. Tversky, D. Kahneman, and A. Schwartz. 1997. The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test. *Quarterly Journal of Economics* 112(2): 647–61.

Thibaut, J. W., and H. H. Kelley. 1959. *The Social Psychology of Groups*. New York: Wiley.

Thorndike, E. L. 1898. Animal Intelligence: An Experimental Study of the Associative Processes in Animals. *Psychological Review Monograph Supplement* 2.

———. 1911. *Animal Intelligence*. New York: Thoemmes Press Classics in Psychology (1998).

Tirole, J. 1996. A Theory of Collective Reputations (with Applications to the Persistence of Corruption and to Firm Quality). *Review of Economic Studies* 63: 1–22.

Tolman, E. C. 1925. Purpose and Cognition: The Determinants of Animal Learning. *Psychological Review* 32: 285–97.

Twain, Mark. 1897. *Following the Equator—A Journey Around the World.* Hartford, CT: American Publishing Co.

Ungemach, C., N. Chater, and N. Stewart. Forthcoming. Decisions from Experience without Sampling Error.

Van Boening, M., A. W. Williams, and S. LaMaster. 1993. Price Bubbles and Crashes in Experimental Call Markers. *Economic Letters* 41: 179–85.

Van Huyck, J. B., R. C. Battalio, and R. O. Beil. 1990. Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure. *The American Economic Review* 80: 234–48.

———. 1991. Strategic Uncertainty, Equilibrium Selection, and Coordination Failure in Average Opinion Games. *The Quarterly Journal of Economics* 106: 885–911.

Van Huyck, J. B., R. C. Battalio, and F. Rankin. 2007. Selection Dynamics and Adaptive Behavior without Much Information. *Economic Theory* 33(1): 53–65.

Van Huyck, J. B., J. Cook, and R. Battalio. 1997. Adaptive Behavior and Coordination Failure. *Journal of Economic Behavior and Organization* 32(4): 483–503.

Vega-Redondo, F. 1997. The Evolution of Walrasian Behavior. *Econometric* 65(2): 375–84.

Viscusi, W. K. 2002. *Smoke-Filled Rooms: A Postmortem on the TobaccoDeal*. Chicago: University of Chicago Press.

von Neumann, J., and O. Morgenstern. 1947. Theory of Games and Economic Behavior, 2nd ed. Princeton, NJ: Princeton University Press.

Vulkan, N. 2000. An Economist's Perspective on Probability Matching. *Journal of Economic Surveys* 14: 101–18.

Wakker, P. 2010. *Prospect Theory: For Risk and Ambiguity*. Cambridge, UK: Cambridge University Press.

Walker, M., and J. Wooders. 2001. Minimax Play at Wimbledon. *American Economic Review* 91(5): 1521–38.

Waller, C. J., T. Verdier, and R. Gardner. 2002. Corruption: Top Down or Bottom Up? *Economic Inquiry* 40(4): 688–703.

Weber E. U., S. Shafir, and A. Blais. 2004. Predicting Risk Sensitivity in Humans and Lower Animals: Risk as Variance or Coefficient of Variation. *Psychological Review* 111(2): 430–45.

Weber, R. A. 2003. Learning with No Feedback in a Competitive Guessing Game. *Games and Economic Behavior* 44: 134–44.

Wegner, D., D. Schneider, S. Carter, and T. White. 1987. Paradoxical Effects of Thought Suppression. *Journal of Personality and Social Psychology* 53(1): 5–13.

Wilcox, N. 2006. Theories of Learning in Games and Heterogeneity Bias. *Econometrica* 74: 1271–92.

Yechiam, E., G. Barron, and I. Erev. 2005. The Role of Personal Experience in Contributing to Different Patterns of Response to Rare Terrorist Attacks. *Journal of Conflict Resolution* 49: 430–39.

Yechiam, E., and J. R. Busemeyer. 2006. The Effect of Foregone Payoffs on Underweighting Small Probability Events. *Journal of Behavioral Decision Making* 19: 1–16.

Yechiam, E., J. R. Busemeyer, J. C. Stout, and A. Bechara. 2005. Using Cognitive Models to Map Relations between Neuropsychological Fisorders and Human Decision Making Deficits. *Psychological Science* 16(12): 973–78.

Yechiam, E., I. Erev, and G. Barron. 2006. The Effect of Experience on Using a Safety Device. *Safety Science* 44: 515–22.

Yechiam, E., I. Erev, and D. Gopher. 2001. On the Potential Value and Limitations of Emphasis Change and Other Exploration-Enhancing Training Methods. *Journal of Experimental Psychology: Applied* 7(4): 277–85.

Yechiam, E., J. Kanz, A. Bechara, J. C. Stout, J. R. Busemeyer, E. M. Altmaier, and J. Paulsen. 2008. Neurocognitive Deficits Related to Poor Decision-Making in People Behind Bars. *Psychonomic Bulletin and Review* 15: 44–51.

Zimbardo, P. G. 1969. The Human Choice: Individuation, Reason, and Order versus Deindividuation, Impulse, and Chaos. *Nebraska Symposium on Motivation* 17: 237–307.

Zohar, D. 1980. Safety Climate in Industrial Organizations: Theoretical and Applied Implications. *Journal of Applied Psychology* 65: 96–102.

Zohar, D., and G. Luria. 2005. A Multilevel Model of Safety Climate: Cross-Level Relationships between Organization and Group-Level Climates. *Journal of Applied Psychology* 90: 616–28.