Journal of Economic Perspectives—Volume 11, Number 4—Fall 1997—Pages 167–188

Progress in Behavioral Game Theory

Colin F. Camerer

Is game theory meant to describe actual choices by people and institutions or not? It is remarkable how much game theory has been done while largely ignoring this question. The seminal book by von Neumann and Morgenstern, *The Theory of Games and Economic Behavior*, was clearly about how rational players would play against others they knew were rational. In more recent work, game theorists are not always explicit about what they aim to describe or advise. At one extreme, highly mathematical analyses have proposed rationality requirements that people and firms are probably not smart enough to satisfy in everyday decisions. At the other extreme, adaptive and evolutionary approaches use very simple models—mostly developed to describe nonhuman animals—in which players may not realize they are playing a game at all. When game theory does aim to describe behavior, it often proceeds with a disturbingly low ratio of careful observation to theorizing.

This paper describes an approach called "behavioral game theory," which aims to describe actual behavior, is driven by empirical observation (mostly experiments), and charts a middle course between over-rational equilibrium analyses and under-rational adaptive analyses.

The recipe for behavioral game theory I will describe[1] has three steps: start with a game or naturally occurring situation in which standard game theory makes a bold prediction based on one or two crucial principles; if behavior differs from

[1] Interested readers familiar with game theory should read Crawford (1997), whose approach is more eclectic than mine. He concludes (p. 236) that "most strategic behavior can be understood via a synthesis that combines elements from each of the leading theoretical frameworks [traditional game theory, evolutionary game theory, and adaptive learning] with a modicum of empirical information about behavior . . ." Behavioral game theory adds psychological interpretations to this synthesis.

■ *Colin F. Camerer is Rea and Lela G. Axline Professor of Business Economics, California Institute of Technology, Pasadena, California. His e-mail address is ⟨camerer@hss.caltech.edu⟩.*

the prediction, think of plausible explanations for what is observed; and extend formal game theory to incorporate these explanations. This paper considers three categories of modelling principles and catalogues violations of these principles. The first section will focus on cases in which players, rather than focusing self-interestedly on their own payoff alone, seem to respond in terms related to social utility, showing concerns about fairness and the perceived intentions of other players. The next section will focus on problems of choice and judgment: cases in which players respond to differences in how the game is described, rather than to the outcomes, and in which players systematically overestimate their own capabilities. A third section will investigate some elements that people bring to strategic situations that are usually unaccounted for in game theory: a common awareness of certain focal points for agreement, a belief that timing of a choice may confer privileged status or change players' thinking, and a natural instinct to look only one or two levels into problems that permit many levels of iterated reasoning.

Organizing findings in this way is like taking a car engine apart and spreading out the parts, so that each part can be inspected separately and broken ones replaced. The hope is that working parts and necessary replacements can later be reassembled into coherent theory. Just as the rebuilt car engine should run better than before, the eventual goal in behavioral game theory is to be able to take a description of a strategic situation and predict actual behavior at least as well as current theories do. Better descriptive principles should also improve the prescriptive value of game theory, since players who desire to play rationally need to know how others are likely to play.

Simple games are also useful for establishing phenomena that should be incorporated into economics beyond game theory. In experiments, people routinely reject profitable bargains they think are unfair, contribute substantially to public goods and do not take full advantage of others when they can (exhibiting surprisingly little "moral hazard"). Textbook discussions of wage setting, public goods problems and the need for incentive contracts and monitoring to prevent moral hazard paint a different picture, portraying people as more socially isolated, uncooperative and opportunistic than they are in experiments. If the generality of the experimental results is questioned, the generality of the textbook caricature should be, too. Other game experiments show that players will behave "irrationally" when they expect others to behave even more irrationally, which is one common explanation for excessive volatility in financial markets. Establishing and dissecting such effects in games could help inform theorizing about similar behavior in markets.

Games as Social Allocations

Games give payoffs to more than one person. If players care about the financial payoffs of others, the simplifying assumption of pure self-interest, common to so much game theory analysis, must be modified. Most theories sidestep this concern by assuming the individual utilities already take into account the possibility that players care about how much others get, so that players still want to maximize their own individual utilities. But in practice, when payoffs are measured in dollars or other numerical units, predictions depend on the precise form of the "social utility" function that players use to combine their own payoffs and payoffs of others to decide their preferences.

Simple bargaining games have proved to be useful tools for bringing out issues of social utility. For example, in an ultimatum game, a Proposer offers a division of a sum of money X to a Responder who can accept or reject it. If the Responder accepts the offer, then both players receive the amount of money given; if the Responder rejects the offer, then both players receive nothing. The ultimatum game is an instrument for asking Responders, "Is this offer fair?" It forces them to put their money where their mouth is and reject offers they claim are unfair. Dozens of studies with these games establish that people dislike being treated unfairly, and reject offers of less than 20 percent of X about half the time, even though they end up receiving nothing. Proposers seem to anticipate this behavior, and to reduce the risk of rejection, they typically offer 40–50 percent of X. The basic result has been replicated in several countries for stakes of up to two months' wages.

The extent of pure altruism by Proposers can be measured with a "dictator" game, an ultimatum game in which the Proposer dictates the division of the money, because the Responder cannot reject the offer. In the dictator game, Proposers offer substantially less than in ultimatum games, but still generally offer an average of 20–30 percent of the sum to be divided.[2] Recent experiments on "trust" (or "gift exchange") add a stage in which one player can either keep a fixed sum or give a larger sum to a second player, who then allocates it between the two players however she likes, as in a dictator game. The first player can either keep the fixed sum, or trust the second player to give back more. In these games, there is often a surprising amount of trust, and the "return to trust" is slightly positive (Fehr, Kirchsteiger and Reidl, 1993; Berg, Dickhaut and McCabe, 1995; and Bolle, 1995).

The crucial step, of course, is to incorporate findings like these into a theory that is more general but still reasonably parsimonious. A natural alternative is to assume that player 1's utility function incorporates the comparison between 1's payoff and player 2's in some way. For example, the separable form in which the utility of player 1 over the consumption of both players, x_1 and x_2, is given by $u_1(x_1,x_2) = v(x_1) + \alpha v(x_2)$ allows for both sympathy or altruism. Sympathy occurs when $\alpha > 0$, so that player 1 benefits from the consumption of player 2; envy occurs when $\alpha < 0$, so that player 1 suffers from the consumption of player 2. "Sympathy coefficients" were mentioned by Adam Smith, and the linear formulation was discussed by Edgeworth (1881 [1967], pp. 101–102), who wrote: "We must modify the utilitarian integral . . . by multiplying each pleasure, except the pleasure of the

[2] For a brief introduction to the literature on ultimatum and dictator games in this journal, see the "Anomalies" column by Camerer and Thaler in the Spring 1995 issue.

Table 1
A Prisoner's Dilemma

	Cooperate	Defect
Cooperate	3, 3	0, 5
Defect	5, 0	1, 1

agent himself, by a fraction—a factor doubtless diminishing with what may be called the social distance between the individual agent and those of whose pleasures he takes account."

But this sort of model looks only at the final allocation between players, and thus does not accommodate the fact that the way in which an unequal allocation came about, and what that implies to one player about the intentions of the other, also affect behavior. For example, Blount (1995) found that players were more willing to accept uneven offers generated by a chance device than the same offer generated by a Proposer who benefits from the unevenness. Exploring the sources of altruism or envy requires a model in which social values are triggered by actions and intentions of others. Rabin (1993) proposed an elegant model that does so. He suggests (roughly speaking) that player 1 has a positive sympathy coefficient $\alpha > 0$ when player 2 "kindly" helps 1; and conversely, $\alpha < 0$ when player 2 behaves "meanly" by choosing an action that hurts 1. Rabin assumes these feelings add to utility from money payoffs, but become relatively less important as money payoffs rise. These assumptions and a few others lead to "fairness equilibrium." The concept of a fairness equilibrium is consistent with observations from three games that are anomalies for most other approaches.

In an ultimatum game, an unfair offer by the Proposer is mean, and hence triggers reciprocated meanness from the Responder (a rejection). An uneven offer generated by chance, however, is not a mean act by the player who benefits from it, so fairness equilibrium predicts correctly that such offers are rejected less often.

In a prisoner's dilemma game, illustrated in Table 1, a rational player recognizes that regardless of whether the other player defects or cooperates, the rational act is to defect. However, when both players follow this logic, they end up worse off than if they could have agreed to cooperate. In experimental games, players seem unexpectedly skillful at avoiding the dilemma and finding their way to cooperation. The concept of fairness equilibrium helps to explain why. At least for small stakes, cooperation in a prisoner's dilemma is a fairness equilibrium. After all, cooperating means sacrificing to help another person, which triggers a reciprocal preference to cooperate. Cooperation is emotionally strategic in this approach, transforming the prisoner's dilemma into a coordination game in which players desire to coordinate their levels of niceness. This jibes with the widely observed fact that players who expect others to cooperate are more likely to cooperate themselves (Sally, 1995).

Table 2
Chicken

		Player 2 Dare (D)	Chicken (C)
Player 1	Dare (D)	$-4x$, $-4x$	$3x$, 0
	Chicken (C)	0, $3x$	x, x

But intentions matter. Suppose another player is forced to cooperate—perhaps by a game structure in which that person cannot choose to defect. Their forced cooperation does not give the other player an unusually high payoff (it is not "nice," because the forced player did not have the option to defect and treat the other player badly). So fairness equilibrium predicts the free-to-choose player is under no obligation to reciprocate and will be more likely to defect than in the standard dilemma.

The game of chicken, it turns out, is perhaps the ideal game for contrasting fairness and self-interested preferences. Table 2 gives the payoffs in "chicken." In this game, both players would like to Dare (D) the other to Chicken out (C) (then D earns $3x$, and C earns 0) but if both Dare they each earn $-4x$. The players move simultaneously. The (pure strategy) Nash equilibria are (D,C) and (C,D), since a player who expects the other to Dare should Chicken out, and vice versa. However, fairness equilibrium predicts exactly the opposite, at least for small stakes. Start in the upper right (D,C) cell. Though the game is actually played simultaneously, suppose the players' reason in their own minds about moves and countermoves, in a kind of mental tatonnement, before deciding what to do. If player 1 moves from this cell, politely choosing C, she sacrifices $2x$ (getting x instead of $3x$) to benefit player 2 by the amount x. This nice choice triggers reciprocal niceness in player 2; rather than exploiting player 1's choice of C by responding with D, player 1 prefers to sacrifice (settling for x instead of $3x$) to "repay" player 1's kindness. Thus, both politely playing Chicken is a fairness equilibrium. By opposite reasoning, (D,D) is a mean fairness equilibrium; rather than back down in the face of the other's D, both would rather lose more by picking D, to hurt their enemy.[3] In a recent study of chicken, 60 percent of the observations in the last half of the experiment were fairness equilibrium choices (C,C) and (D,D), and only 12 percent were Nash equilibria (D,C) and (C,D) (McDaniel, Rutström and Williams, 1994).

[3] The "mean" fairness equilibrium (D,D) illustrates the advantage of chicken over the prisoner's dilemma for studying social values. In the prisoner's dilemma, defection is both a self-interested choice (reflecting neither niceness nor meanness), *and* it is the choice a mean-spirited (or envious) person would make. In chicken, the best response of a self-interested person to an expectation that the other person would play D is C, but a mean person would pick D.

The fairness equilibrium or "reciprocated value" model is a solid new plateau for understanding departures from self-interest in games. The model captures basic facts that the simpler separable and comparative models do not capture, particularly the reciprocal nature of social values and the distinction between uneven outcomes and unfair actions. Its formal specification connects fairness equilibrium closely to standard game theory. Games like chicken allow both nice and mean outcomes to arise in the same game, capturing phenomena like the blissful happiness of a loving couple and their bitter, mutually destructive breakup.

Games Require Choice and Judgment

Rational players will perceive a game and themselves clearly and consistently. However, when "framing effects" are important, players see the game differently according to how it is described. When players are overconfident of their own abilities, they fail in seeing the likely consequences of their actions. This section considers these two phenomena in turn.

Framing Effects

Theories of choice often invoke an axiom of "description invariance," which holds that differences in descriptions that do not alter the actual choices should not alter behavior. A "framing effect" occurs when a difference in description does cause behavior to vary.

For example, give subjects $10 in advance, then ask them whether they would choose a certain loss of $5 (for a net gain of $5) or flip a coin and lose either $10 or 0, depending on the outcome. Those subjects choose to gamble more frequently than subjects who are given nothing and asked to choose between gaining $5 or flipping a coin with $10 and 0 outcomes (Tversky and Kahneman, 1992). Generally, people are more likely to take risks when outcomes are described as losses than when the same outcomes are described as gains. Players in games can exhibit a version of this "reflection effect." Players are more willing to risk disagreement when bargaining over possible losses than when bargaining over possible gains (Neale and Bazerman, 1985; Camerer et al., 1993). In certain coordination games with multiple equilibria, avoidance of losses acts as a "focal principle" that leads players to coordinate their expectations on those equilibria in which nobody loses money (Cachon and Camerer, 1996).

Overconfidence about Relative Skill

The now-standard approach to games of imperfect information pioneered by John Harsanyi presumes that players begin with a "common prior" probability distribution over any chance outcomes. As an example, consider two firms *A* and *B*, who are debating whether to enter a new industry like Internet software. Suppose it is common knowledge that only one firm will survive—the firm with more skilled managers, say—so firms judge the chance that their managers are the more skilled.

The common prior assumption insists both firms cannot think they are each more likely to have the most skill. Put more formally, a game like this can be modelled as a tree where the top node separates the game into two halves—a left half in which A is truly more skilled, and a right half in which B is truly more skilled. The firms can't play coherently if A and B believe they are actually on different halves of the tree. In this way, overconfidence about relative skill violates the common prior assumption.

Of course, this requirement of a common prior does not rule out that players may have private information. In the example of the two competing software firms, each could know about its own secret projects or the tastes of its customers, but others must know what that information could possibly be.

Dozens of studies show that people generally overrate the chance of good events, underrate the chance of bad events and are generally overconfident about their relative skill or prospects. For example, 90 percent of American drivers in one study thought they ranked in the top half of their demographic group in driving skill (Svenson, 1981). Feedback does not necessarily dampen overconfidence much (and could make it worse): one study even found overconfidence among drivers surveyed in the hospital after suffering bad car accidents (Preston and Harris, 1965).

But if those involved in game-like interactions are overconfident, the result may matter, dramatically. For example, economic actors behaving in a mutually overconfident way may invest the wrong amount in R&D, prolong strikes or delay agreements inefficiently, opt for high-risk sports or entertainment careers instead of going to college, and so on.

Although overconfidence has been largely ignored in theorizing about games, there are some clear experimental examples of its effects. In the Winter 1997 issue of this journal, Babcock and Loewenstein review several such examples and also describe what they call "self-serving bias."

One possible economic manifestation of overconfidence is the high failure rates of new businesses (around 80 percent fail in their first three years). Of course, high failure rates are not necessarily inconsistent with profit maximization. Maybe new business owners judge their relative skill accurately, but business returns are positively skewed "lottery ticket" payoffs in which the few survivors are extremely profitable. Then a large percentage might fail, even though the expected value of entering is positive. The overconfidence and rational entry explanations are very difficult to distinguish using naturally occurring data. But they can be compared in an experimental paradigm first described in Kahneman (1988) and extended by Rapoport et al. (forthcoming).

In the entry game paradigm, each of N subjects can choose to enter a market with capacity C or can stay out and earn nothing. The profit for each entrant is the same, but more entrants means that everyone earns a lower level of profit. If C or fewer enter, the entrants all earn a positive profit. If more than C enter, the entrants all lose money. In pure-strategy Nash equilibria, players should somehow coordinate their choices so that exactly C enter and $N - C$ stay out.

Rapoport et al. (forthcoming) found that when subjects played repeatedly with feedback about the number of entrants in each period, about C subjects did enter, even though players could not communicate about which of them would enter and which would stay out.

When all entrants earn the same profit, there is no such thing as a more or less successful entrant, so overconfidence about relative success cannot fuel excessive entry. Dan Lovallo and I (1996) enriched the paradigm to allow that possibility. In one version of the experiment, we informed 14 subjects that they were going to take a trivia quiz and that the six entrants who turned out to be top-ranked in that quiz would share $50 in profits. (The trivia quiz can be considered in economic terms as a difference in product cost or R&D effectiveness that would make certain firms more profitable if they choose to enter.) However, any entrants below the top six would lose $10 each. Notice in this example that if 11 subjects enter, then the total industry profit is zero ($50 divided among the six high scores, and losses of $10 apiece for the other five). Thus, if subjects are risk neutral and think they are equally likely to rank high and low relative to other entrants, about 11 of the 14 players should enter and the rest should stay out. In different versions of the experiment, the number of top-ranked entrants who would receive a share of the positive profits was then systematically varied from six.

We found that in our baseline condition, when the rankings of potential entrants were determined randomly, the number of entrants was typically one or two fewer than the number that would drive industry profits to zero, so the average entrant made a slight profit. However, when the subjects were told that the rankings would be determined by the trivia contest, the number of entrants was typically one or two above the number that would drive industry profits to zero, so the average entrant made losses. The difference in trivia contest entry rates and the random-rank baseline condition was statistically significant. Perhaps even more interesting, when subjects were asked to forecast how many others would enter after the trivia game, they accurately forecast that there would be too many entrants and that the average entrant would lose money—but entrants all thought *they* were above average and would earn money.

The overconfidence bias results from a conflict between wanting to be realistic and wanting to feel good about oneself, and from the psychological "availability" of memories that support the rosy view of oneself. Researchers have documented a number of other systematic biases in probability judgments, including biases resulting from using shortcuts that make difficult judgments easier.

One such shortcut is called "representativeness." People who use this heuristic judge how well a sample represents a statistical process (or a person or category) and use that judgment to estimate the sample's likelihood. This shortcut is sensible, but sometimes conflicts with normative principles of probability. An example from game theory comes from "weak link" coordination games (Van Huyck, Battalio and Beil, 1990). In these games, several players choose numbers from 1 to 7 at the same time. Each player earns a payoff that depends on the minimum number anybody picks—a high minimum is better—and players are penalized for picking a number above the minimum. Because players are penalized for picking too high a number, they would like to match whatever they expect the minimum to be. A little statistical thought suggests that the minimum is likely to be much lower in large groups than it would be in small groups. And it is: in groups of six or more the minimum is usually 1–2, but in groups of two or three the minimum is usually higher. Surprisingly, however, the spread of numbers subjects choose in the first period does not depend on the size of their group. It seem as if players construct a guess about what the "representative" other player will do, then "clone" the representative player several times to represent the group, so they can figure out what the whole group will do. If the representative player's choice is a single number, players who reason this way will not realize intuitively that the minimum number picked by a large group of players will be lower than the minimum from a smaller group. As a result, players in large groups mistakenly pick numbers that are too high and are penalized as a result.

Games as Strategic Situations

Many principles of strategic reasoning that are widely used in game theory have been questioned descriptively. Some of the principles are subtle and took decades for sophisticated theorists to discover and codify, so it seems unlikely that they are applied precisely by average folks. Attention has turned, properly, to the conditions under which principles of strategic reasoning might be learned and what people are thinking before and while they learn. I will focus here on three such principles: irrelevance of labels and timing, iterated dominance and backward induction.

Focal Points and the Irrelevance of Timing

To simplify analyses, game theorists often assume that some features of the game description—like the way strategies are labelled, or the timing of moves—are irrelevant for determining equilibria. But sometimes these features do affect choices.

As one example of how labels can make a difference, Schelling (1960) argued long ago that the labels can create psychologically prominent "focal points" that resolve coordination problems. Schelling gave examples of "matching games" in which players earn a prize if they choose the same strategy as another player, but otherwise get nothing. In matching games, players want to coordinate their choices on some strategy, and they don't care which one it is. Mehta, Starmer and Sugden (1994) report interesting data from matching games. When two strategies are labelled "heads" or "tails," 87 percent of subjects chose "heads," and only 13 percent chose tails. Good coordination can occur even when strategy sets are large. Two-thirds choose "rose" when asked to match a flower name, 59 percent choose "red" from the set of colors, 50 percent choose the boy's name John, and 40 percent choose the number 1 from the (infinite!) set of numbers. Subjects are

Table 3
Battle of the Sexes

		Player Column A	Player Column B	Choice Frequencies Simultaneous	Choice Frequencies Sequential
Player Row	A	0, 0	2, 6	38%	12%
Player Row	B	6, 2	0, 0	62%	88%
Simultaneous		35%	65%		
Sequential		70%	30%		

clearly exhibiting some strategic sophistication, because the frequency of the most common choice is much lower when players are asked only to express preferences, rather than match. For example, when asked to name a favorite day of the year, 88 subjects picked a total of 75 different dates; Christmas was the most popular at 6 percent. But when trying to match with others, 44 percent picked Christmas.

Game theory has a lot to learn from subjects about focal principles, but serious theoretical attention to the topic has been rare. Crawford and Haller (1990) show how focal precedents can emerge over time when games are repeated and players are eager to coordinate. Static theories like Bacharach and Bernasconi (1997) do not explain how focal points come about, but they capture the tradeoff between the chance that people commonly recognize a strategy's distinguishing features and the number of other strategies that share those features. Focal principles have potentially wide economic applications in implicit contracting, evolution of convention, social norms and folk law, and corporate culture (Kreps, 1990).

Timing is another descriptive feature of a game that is often assumed to be irrelevant, but can matter empirically. In laying the foundations of game theory, von Neumann and Morgenstern (1944) deliberately emphasized the central role of information at the expense of timing. They believed that information was more fundamental than timing, because knowing what your opponents did necessarily implies that they moved earlier. Alternatively, if you don't know what your opponents did, you won't care whether they already did it, or are doing it now. Combining these principles implies that information is important but that timing, per se, is not.

But empirical work has found surprising effects of move order in games (holding information constant). Take the battle of the sexes game in Table 3. In this game, one player prefers choice *A* and the other choice *B*, but both players would rather coordinate their choices than end up apart. This game has two pure-strategy multiple equilibria (*A*,*B*) and (*B*,*A*) that benefit players differently. There is also a mixed-strategy equilibrium—choosing *B* 75 percent of the time—that yields an expected payoff of 1.5 to both players.[4] Notice that both players prefer either one of the two pure strategy equilibria to mixing, but they each prefer a different one.[5] Cooper et al. (1993) found that when players move simultaneously, they converge roughly to the mixed strategy equilibrium, choosing *B* more than 60 percent of the time, as shown in Table 3. In a sequential condition, say that the Row player moves first, but her move is not known to Column. In this case, Row players choose their preferred equilibrium strategy, *B*, 88 percent of the time, and Column players go along, choosing *A* 70 percent of the time. The mere knowledge that one player moved first, without knowing precisely how she moved, is enough to convey a remarkable first-mover advantage that the second-mover respects. The data suggest a magical "virtual observability," in which simply knowing that others have moved earlier is cognitively similar to having observed what they did (Camerer, Knez and Weber, 1996). After all, if Column figures out that Row probably selected *B*, Column's sensible choice (setting aside mean retaliation) is to go along.

It is not clear how virtual observability works, but it appears that when one player explicitly moves first, other players think about the first-mover's motivations more carefully. If first-movers anticipate this, they can choose the move that is best for themselves, because they know players moving later will figure it out. Psychology experiments have established related ways in which reasoning about events depends curiously on their timing. For example, many people dislike watching taped sports events. Even when they don't know the outcome, simply knowing that the game is over drains it of suspense. People can also generate more explanations for an event that has already happened than for one that has yet to happen.

One experiment investigated the psychology of timing in the game of "matching pennies" (Camerer and Karjalainen, 1992), in which both players independently choose heads (*H*) or tails (*T*). In this game, one player wants to match, but the other player wants to mismatch. If the "mismatching" player moves first, she is more likely to choose either *H* or *T*, trying to outguess what the first mover will later do than to choose a chance device which explicitly randomizes between *H* and *T* for her. But if the "matching" player has already moved, the mismatching player is more likely to choose the chance randomizing device, hedging her bet. Apparently people are more reluctant to bet on their guesses

[4] Intuitively, think of the mixed strategy equilibrium in this way. If player 1 knows that player 2 will choose *B* 75 percent of the time, then for player 1, the expected value of choosing *A* will be $.25(0) + 2(.75) = 1.5$, and the expected value of choosing *B* will be $(.25)6 + .75(0) = 1.5$. In other words, the highest payoff for player 1 is 1.5. Of course, the same logic works in reverse; if player 2 knows that player 1 will choose *B* 75 percent of the time, then the highest possible payoff for player 2 is 1.5. Therefore, if both players choose *B* 75 percent of the time, then the best response of both players to that choice will involve a payoff of 1.5.

[5] The outcome (*B*,*B*) is also a "mean" fairness equilibrium. Suppose that player 1 thinks that player 2 expects them to play *B*, and as a result, player 2 is going to respond meanly by choosing *B* to harm player 1 (and themselves). Then, a mean-spirited player 1 will choose *B* in a sort of preemptive retaliation, so that a mutually destructive (*B*,*B*) equilibrium results.

about what other players have already done than on guesses about what other players will later do.

Iterated Dominance and "Beauty Contests"

"Iterated dominance" is the strategic principle that means that players first rule out play of dominated strategies by all players, then eliminate strategies that became dominated after the first set was eliminated, and so forth. In many games, this iterative process yields a unique choice after enough steps of iterated dominance are applied.[6]

But do people actually apply many levels of iterated dominance? There are many reasons for doubt. Studies of children show that the concept "beliefs of others" develops slowly. Psycholinguist Herb Clark studies how people infer the meaning of statements with vague references ("Did he do it already?"), which require people to know what others know, what others know they know, and so forth. Clark jokes that the grasp of three or more levels of iterated reasoning "can be obliterated by one glass of decent sherry." Since the process of iteration depends on beliefs about how others will play (and their beliefs . . .), then if even a few people behave irrationally, rational players should be cautious in applying iterated dominance.

Experiments are useful for measuring where the hierarchy of iterated dominance reasoning breaks down. An ideal tool is the "beauty contest game," first studied experimentally by Nagel (1995). A typical beauty contest game has three rules. First, N players choose numbers x_i in [0,100]. Second, an average of the numbers is taken. Third, a target is selected that is equal to the average divided by the highest possible choice. For illustration, say the average number chosen is 70, so the target is 70 percent of the average number. Finally, the player whose number is closest to the target wins a fixed prize. (Ties are broken randomly.) Before proceeding, readers should think of what number they would pick if they were playing against a group of students.

This game is called a "beauty contest" after the famous passage in Keynes's (1936, p. 156) *General Theory of Employment, Interest, and Money* about a newspaper contest in which people guess what faces others will guess are most beautiful. Keynes used this as an analogy to stock market investment. Like people choosing the prettiest picture, players in the beauty contest game must guess what average number others will pick, then pick 70 percent of that average, while knowing that everyone is doing the same.

The beauty contest game can be used to distinguish the number of steps of reasoning people are using. Here's how: suppose a player understands the game and realizes that the rules imply the target will never be above 70. To put it another way, numbers in the range [70,100] violate first-order iterated dominance. Now suppose a subject chooses below 70 and thinks everyone else will as well. Then the

[6] These "dominance solvable" games include Cournot duopoly, finitely repeated prisoner's dilemma and some games with strategic complementarities.

Table 4
Beauty Contest Results
(Singapore 7-person groups, p = 7)

Round	*Mean*	*Median*	*Std. Dev.*	*Percentage Choosing 0*
1	46.07	50	28.04	0.02
2	31.20	28	17.69	0.03
3	25.47	20	20.57	0.00
4	18.79	15	16.23	0.00
5	18.55	10	22.53	0.00
6	15.29	10	18.00	0.05
7	16.31	10	21.53	0.09
8	14.85	8	20.30	0.12
9	15.36	7	22.51	0.11
10	13.89	6	22.53	0.19

subject can infer that the target will be below .7 × 70, or 49, so an optimal choice is in the range [0,49]. Hence, a choice between [49,70] is consistent with a subject being first-order rational, but not being sure others are rational. The next range of choices, [34.3,49], is consistent with second-order rationality but violates third-order rationality. In this way, number choices in the beauty contest game reveal the level of iterated rationality. Infinitely many steps of iterated dominance leads to the unique Nash equilibrium—pick zero.

Table 4 summarizes results from an experimental study with a beauty contest game that involved predicting what 70 percent of the average choice would be (Ho, Camerer and Weigelt, forthcoming). The game was carried on for 10 rounds, with subjects receiving feedback after each round. First-round choices were typically dispersed around 31–40. Few subjects chose the equilibrium of zero in the first round—nor should they have! Some subjects violated dominance by choosing large numbers near 100, but not very many. As the rounds progress, choices are drawn toward zero as subjects learn. Thus, some notion of limited iterated reasoning is essential to understand initial choices and the movement across rounds. Econometric analysis indicates that most players look one or two iterations ahead (Holt, 1993; Stahl and Wilson, 1995).

Since the beauty contest game is easy to conduct, informative and fun, I have collected data from many subject pools playing once for a $20 prize. The mean, median and standard deviation of choices from several groups are shown in Table 5. Some are highly educated professionals—economics Ph.D.'s, portfolio managers and the Caltech Board of Trustees, which includes a subsample of 20 CEOs, corporate presidents and board chairmen. Others are college students from three continents and Los Angeles high school students. The most remarkable fact is that the average choices are very similar for all these groups. Even in the highly educated groups, only 10 percent of the subjects choose 0 (and don't win!). Quadrupling the prize brings numbers only slightly closer to the equilibrium.

Table 5
Beauty Contest Results from Many Subject Pools

	Mean	*Median*	*Std. Dev.*	*Percentage Choosing 0*	*Sample Size*
Portfolio Managers	24.31	24.35	16.15	0.08	26
Economics Ph.D.'s	27.44	30.00	18.69	0.13	16
Caltech Board of Trustees					
All	42.62	40.00	23.38	0.03	73
CEOs only	37.81	36.50	18.92	0.10	20
College Students					
Caltech	21.88	23.00	10.35	0.07	27
Germany	36.73	33.00	20.21	0.03	67
Singapore	46.07	50.00	28.04	0.02	98
UCLA	42.26	40.50	17.95	0.00	28
Wharton	37.92	35.00	18.84	0.00	35
High school students (U.S.)	32.45	28.00	18.61	0.04	52

Of course, specially trained players might choose closer to the equilibrium,[7] or experts hired to advise institutions might "figure out" that their clients should play zero. But the experts would only be right if other players actually do choose zero. The trick is to be one step of reasoning ahead of the average player, but no further![8]

These results suggest that instead of assuming that players apply iterated rationality through many levels, it is more realistic to assume a limited number of iterations, perhaps with the number of levels depending on the characteristics of the subjects involved. Furthermore, players who use equilibrium analysis alone to guide their choices are making a mistake. Anticipating that other players use limited iterated reasoning, or that strategy labels or timing create focal points that others will realize and act upon, leads players to behave more intelligently.

Backward Induction and Subgame Perfection

A central concept in game theory is "subgame perfection." An equilibrium pattern of behavior is subgame perfect if players think about every possible "subgame" that could be reached later in a game tree, guess what players would do in those subgames, and use the guesses in deciding what to do at the start. Implementing this process requires players to reason about future events and "backward induct" to the present. Subgame perfection is important because games

[7] However, keep in mind that Caltech undergraduates are much more analytically capable than average. Last year, the median math SAT score among entering students was 800.

[8] Some research on leadership, reflecting the same principle, suggests the ideal leader should be one standard deviation more intelligent than the group he or she leads, but no smarter than that.

often have several equilibria, and those that are not subgame perfect are considered to be less likely to occur.

As a reasoning principle, however, backward induction is descriptively dubious because studies of how people learn to play chess and write computer programs show that backward reasoning is unnatural and difficult. And backward induction requires players to spend precious time thinking about future events that seem unlikely to occur. Should they bother?

Direct tests of backward induction in games come from work on sequential bargaining (Camerer et al., 1993). In these experiments, player 1 offers a division of a pie. If player 2 accepts the offer, the game ends. But if player 2 rejects the offer, then the pie shrinks in size and player 2 offers a division to player 1. Again, if player 1 accepts the offer then the game ends. Otherwise, the pie shrinks again in size and player 1 again gets to offer a division. If the third-round offer is rejected, the game ends and players get nothing. This is a game of backward induction, where the optimal offer can be reached by working back from the last period. If the third pie is reached, and play is rational, then player 1 will offer only an epsilon slice to player 2, who will accept. Knowing this, player 2 recognizes that when dividing the second pie, she must give player 1 a slice equal to the smallest pie (plus epsilon), and keep the remainder, or else player 1 will reject the offer. Knowing this, player 1 recognizes that when dividing the first pie, offering player 2 a slice equal to the size of the second pie minus the third (plus epsilon), will be an offer that player 2 will accept.

Subjects trained briefly in backward induction reach this result readily enough. But as in many other experiments, first-round offers of untrained subjects lay somewhere between dividing the first pie in half, and the equilibrium offer (pie two minus pie three). More interestingly, the experiment was carried out on computers, so that to discover the exact pie sizes in the three rounds, subjects had to open boxes on a computer screen. Measurements of the cursor's location on the screen indicated the order in which boxes were open and how long they were kept open.[9] By presenting the game to subjects in this way, the subjects are forced to reveal the information they are looking at, giving clues about their mental models and reasoning. Subjects tended to look at the first-round pie first, and longest, before looking ahead, contrary to the "backward induction" looking pattern exhibited by trained subjects. In fact, subjects did not even open the second- and third-round boxes—ignoring the sizes of the second and third pies entirely—on 19 percent and 10 percent of the trials, respectively. These subjects simplify a difficult problem by ignoring future choice nodes that seem unlikely to ever be reached. Their heuristic might be considered sensible, because nearly 90 percent of the trials ended after one round.

[9] Psychologists have used similar methods for nearly 100 years, recording movements of eyes as people read, to understand how people comprehend text.

Speculations and New Directions

Systematic violations of game-theoretic principles are not hard to find because all useful modelling principles are simplifications, and hence are sometimes false. Table 6 summarizes the discussion to this point by offering a list of a few principles that are widely used in game theory, along with the systemic violations of those principles and citations for selected experiments documenting those findings. Lists of this sort are a start. The next step is to use the evidence of violations to construct a formal and coherent theory. Substantial progress has already been made in two areas mentioned earlier: measuring social values and extending game theory to include them; and measuring and incorporating differences among players, like players using different steps of iterated reasoning in beauty contests. Behavioral game theory could usefully extend standard theory in three other ways.

Nonexpected utilities. Players do not always choose the strategy with the highest expected utility. They sometimes value losses differently than gains, and can have aversions toward (or preferences for) strategic ambiguity or uncertainty. Several models have been proposed to bring pattern to this behavior. In the prospect theory of Tversky and Kahneman (1992), people value gains and losses from a reference point (rather than final wealth positions) and dislike losses much more than they like equal-sized gains, which can explain why describing payoffs as gains or losses matters in some experiments. Aversion toward ambiguity can be explained by models that use probabilities that are nonadditive.

A less psychologically grounded approach is to allow the possibility of errors in choice or uncertainty over payoffs that imply that while players are more likely to choose the strategy with the highest expected utility, they are not certain to do so. McKelvey and Palfrey (1995) propose what they call a "quantal response" function, which inserts a variable into the choice function to capture the degree of randomness in decisions. A "quantal response equilibrium" exists if players know how much randomness is in the decisions of others and choose accordingly.[10] This approach offers a parsimonious method of explaining several different behavioral phenomena. For example, if some players do not always make the optimal choice, then other players should use only a limited number of iterated rational steps in situations like the beauty contest games. Or in the ultimatum game, a responder may be likely to reject a smaller offer mistakenly, because it is a small mistake, and knowing this, rational proposers are not likely to make very uneven offers. The quantal response approach also seems to explain subtle experimental patterns in contributions to public goods (Anderson, Goeree and Holt, 1996).

Learning. Much recent interest has been focussed on adaptive learning models,

[10] A handy quantal response function is the logit form, $P(S_i) = e^{\lambda\pi(S_i)}/\Sigma_i e^{\lambda\pi(S_i)}$ (where $\pi(S_i)$ is the expected payoff of strategy S_i). The constant λ captures imprecision in choices, and could be interpreted as either sensitivity to dollar payoffs, or existence of nonpecuniary utilities unobserved by the experimenter. If $\lambda = 0$, then players choose equally often among strategies. As λ grows larger, $P(S_i)$ approaches one for the strategy with the highest expected payoff (the best response).

Table 6
Evidence on Game Theory Modelling Principles

Game-Theoretic Modelling Principle	*Systematic Violation*	*Selected Experimental Evidence*
Social Utility		
Independence of payoff utilities from . . .		
a. Payoffs of others	Altruism, envy (dictator games)	Berg, Dickhaut and McCabe (1995)
b. "Intentions" of others	Reciprocal fairness, (ultimatum, PD, chicken, trust)	Ledyard (1995); McDaniel, Rutström and Williams (1994)
Choice and Judgment		
a. Invariance to game description (no gain-loss asymmetry)	More disagreements over losses	Neale and Bazerman (1985); Camerer et al. (1993)
b. Mutually-consistent beliefs (common prior)	Overconfidence about outcomes, relative ability	Neale and Bazerman (1983); Camerer and Lovallo (1996)
Strategic Principles		
a. Irrelevance of strategy labels & timing	Focal points, "virtual observability" of earlier moves	Schelling (1960); Cooper et al. (1993)
b. Iterated dominance	Limited steps of iterated dominance (1–3)	Nagel (1995); Ho, Camerer and Weigelt (forthcoming)
c. Backward induction	Limited look-ahead	Camerer et al. (1993)

which come in two basic forms. "Belief-based" models presume that subjects form beliefs about what others will do, based on past observations, and choose the strategy that maximizes utility given these beliefs (the "best response") (Crawford, 1995). "Reinforcement" models ignore beliefs and assume that strategies have different probabilities or propensities of being chosen, which change as successful strategies are "reinforced" by observed successes or failures (Roth and Erev, 1995).[11] Both kinds of models are narrow. Belief learners pay no special attention to their payoff history. Reinforcement learners pay no attention to the outcomes of strategies they didn't choose, and they don't keep track of choices by others. Both learners ignore information about other players' payoffs. Teck Ho and I (1997) recently developed a general model that synthesizes the two approaches, thereby avoiding some of the weaknesses

[11] Reinforcement models were popular in cognitive psychology until about 30 years ago, when they were largely replaced by the information processing approach (brains are like computers) and, more recently, by "connectionism" (brains are neural networks). Behaviorism was discredited as a general theory of human learning because it could not easily explain higher-order cognition, like language, and lacked neuroscientific detail. These failures also make it unlikely that reinforcement models can fully explain human learning in strategic situations.

of each. Econometric estimates from five out of six data sets strongly reject the two simpler approaches in favor of the synthetic approach.

"Pregame" theory. Except for the beauty contest game, all the experiments described so far in this paper show the players the game in the form of a matrix or tree. The experiments therefore control for many ways in which subjects might misperceive or simplify a game if it were presented verbally. A useful body of research in this area has been compiled in negotiations research (Neale and Bazerman, 1991). Because those experiments are meant as models of realistic negotiations, not tests of game theory, the subjects usually do not know the precise game associated with the negotiation. These unstructured games are of special use for behavioral game theory, because they allow study of the "mental models" of games people construct. Such experiments are necessary for testing the joint hypothesis that subjects represent the situation correctly (by constructing the proper game theory tree or matrix) and that subjects play according to the axioms of game theory.

As an example, Bazerman and Samuelson (1983) studied a diabolically tricky bidding problem with adverse selection they called "acquire-a-company" (based on Akerlof's (1970) famous "lemons" paper). In this game, the shares of a target firm are worth an amount V uniformly distributed between [\$0, \$100]. The target knows its own value V, but the bidder does not. Because of improved management or corporate synergy, the shares are worth 50 percent more to the acquiring company. The rules are simple: the acquiring company bids a price P, and the target automatically sells its shares if and only if the true value V is less than P. What is the bid that maximizes expected profits?

The typical subject bids between \$50–\$75, a result that has been replicated with many subject pools (including business executives and Caltech undergraduates) and over 20 trials with feedback. The logic behind this bid is presumably that the expected value of the company is \$50, and that with a 50 percent premium, this would be worth \$75 to the acquirer—thus bids ranging from \$50–\$75. But this analysis leaves out the fact that the target will not accept a bid unless it is higher than the actual value. Once this adverse selection is taken into account, the optimal bid is actually zero.

To follow this logic in practice, consider a typical bid of \$60. The target will reject this bid if the true value of the firm exceeds \$60. Therefore, by bidding \$60, the bidder can only win the company if value under present ownership is between \$0 and \$60. Given the even probability distribution, this means that a bid of \$60 will only be accepted when the expected value is \$30. Even with a 50 percent premium for the acquirer, the shares will only be worth \$45 after acquisition. Thus, the act of bidding rules out receiving the company if it is worth more than the bid, and this adverse selection always more than offsets the 50 percent value increment.[12] The acquiring company can only lose (in expected value terms) and therefore should not bid.

[12] The formal analysis is not hard: a bid of P leads to an expected target value, conditional on acceptance of the bid, of $E(V|V \leq P) = P/2$. So the bidder's expected profit is the chance of getting the company times the expected profit if they get it, or $(P/100)[1.5E(V|V \leq P)\text{-}P] = ((3P/4)\text{-}P)P/100$, which is negative for all values of P.

The acquire-a-company problem is amazing because the adverse selection is clear once it is pointed out, but subjects invariably miss it and bid too much. The joint hypothesis that the subjects represent the game correctly and play game theoretically is clearly rejected. A common interpretation (Neale and Bazerman, 1991, p. 73) is that "individuals systematically exclude information from their decision processes that they have the ability to include. . . . Subjects who make this mistake are systematically ignoring the cognitions of the other party."

The first part of this interpretation is surely correct, but the second conclusion is not generally true. In all the experiments reviewed in this paper, for example, subjects exhibit some strategic sophistication, forming a guess about the likely thinking and actions of others: proposers make ultimatum offers that accurately forecast what others will reject; knowing that the other player moved first in a battle-of-the-sexes affects the second move; players in entry games forecast the number of other entrants quite accurately; and most beauty contest players apply one or two steps of iterated dominance.

But if bidders in the acquire-a-company game are not ignoring the cognitions of the target, why do they bid too much? An alternate answer is that subjects simplify the problem; they reduce the uncertainty about the firm's true value by substituting its expected value of $50. Two bits of evidence support this notion. First, consider a version of the game without adverse selection—neither the target nor the acquirer is sure of the true value. Then it is optimal for the bidder to substitute an expected value for the distribution of value, and bid somewhat higher than that, which is precisely what subjects do in experiments. Second, in versions of the game where the target value has only three possible values (like $0, $50 or $100), so that substituting the expected value for all the possible values is not much of a shortcut, nearly half the subjects spontaneously draw a matrix of all possible target values on paper, compute the expected payoffs of bids and deduce the correct answer. Thus, what the acquire-a-company problem really shows is that an oversimplified mental model of a complex game can lead to violations of game theoretic predictions, not because subjects are necessarily incapable of reasoning game theoretically, but because players constructed the wrong model of the game.

Conclusion

Behavioral game theory aims to replace descriptively inaccurate modelling principles with more psychologically reasonable ones, expressed as parsimoniously and formally as possible.

This approach raises the question of whether traditional game theory could still be useful for advising people how to play. The answer depends on the game and whether deviations are a mistake or not. Some patterns observed in experiments can be considered mistakes that could be avoided by doing what game theory prescribes. For example, players who overconfidently enter an industry thinking their prospects are above average should keep in mind that the other entrants think the same, but only some entrants can be right. Players who do not use backward induction can only benefit from doing so (while realizing that others might not be).

In other cases, game theory provides bad advice because underlying assumptions do not describe other players. Knowing how others are likely to deviate will help a player choose more wisely. For example, it is generally dumb to choose the equilibrium of zero in a beauty contest (even playing against CEOs or brilliant Caltech undergraduates). It is smarter to know the number of reasoning steps most people are likely to take and optimize against that number, and to understand the adaptive process that changes others' choices over time. Similarly, players making offers in ultimatum games should know that many players simply regard an offer of 10 percent as unfair and prefer to reject it.

Experiments have supplied tentative answers to some sharply posed questions. Is there a formal way to incorporate reciprocal social values like fairness, altruism, or revenge, which are widely observed in the lab? Yes: Rabin's (1993) fairness equilibrium. Do judgment phenomena like overconfidence about relative skill matter in games? Yes, but we don't yet know how to include these phenomena in formal extensions of game theory. How many steps of iterated dominance do people use? One to three. Do learning and the construction of mental models in unstructured games matter? Yes, but we need much additional data on how they matter. Moving from these sorts of observations to coherent new modelling is the primary challenge for the next wave of research.

A final caveat: The desire to improve descriptive accuracy that guides behavioral game theory does not mean game theory is always wrong. Indeed, it may be only a small exaggeration to conclude that in most games where people gain experience, equilibrium is never reached immediately and always reached eventually. But this is no triumph for game theory until it includes explanations for behavior in early rounds and the processes that produce equilibration in later rounds.

■ *Comments from Linda Babcock, Kong-Pin Chen, Vince Crawford, Bob Gibbons, Teck Ho, Matthew Rabin, JEP editors (Brad De Long, Alan Krueger and Timothy Taylor) and many seminar audiences were helpful. This work resulted from many collaborations and conversations, especially with Gerard Cachon, Teck Ho, Eric Johnson, Risto Karjalainen, Marc Knez, Richard Thaler, Roberto Weber and Keith Weigelt. Support of NSF SBR 9511001 and the Russell Sage Foundation is gratefully appreciated.*

References

Akerlof, George A., "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics,* May 1970, *105,* 255–83.

Anderson, Simon P., Jacob K. Goeree, and Charles A. Holt, "A Theoretical Analysis of Altruism and Decision Error in Public Goods Games," working paper, University of Virginia, Department of Economics, 1996.

Babcock, Linda, and George Loewenstein, "Explaining Bargaining Impasse: The Role of Self-Serving Biases," *Journal of Economic Perspectives,* Winter 1997, *11*:1, 109–26.

Bacharach, Michael, and Michele Bernasconi, "The Variable Frame Theory of Focal Points: An Experimental Study," *Games and Economic Behavior,* April 1997, *19,* 1–45.

Bazerman, Max H., and W. F. Samuelson, "I Won the Auction but Don't Want the Prize," *Journal of Conflict Resolution,* December 1983, *27,* 618–34.

Berg, Joyce, John W. Dickhaut, and Kevin A. McCabe, "Trust, Reciprocity, and Social History," *Games and Economic Behavior,* July 1995, *10,* 122–42.

Bolle, Friedel, "Does Trust Pay?" Diskussionpapier 14/95, Europa-Universität Viadrina Frankfurt, Oder, November 1995.

Blount, Sally, "When Social Outcomes aren't Fair: The Effect of Causal Attributions on Preferences," *Organizational Behavior and Human Decision Processes,* August 1995, *63*:2, 131–44.

Cachon, Gérard, and Colin Camerer, "Loss-Avoidance and Forward Induction in Experimental Coordination Games," *Quarterly Journal of Economics,* February 1996, *111,* 165–94.

Camerer, Colin F., and Teck-Hua Ho, "Experience-Weighted Attraction Learning in Games: A Unifying Approach." California Institute of Technology Working Paper No. 1003, 1997.

Camerer, Colin F., and Risto Karjalainen, "Ambiguity-Aversion and Non-Additive Beliefs in Noncooperative Games: Experimental Evidence." In Munier, B., and M. Machina, eds., *Models and Experiments on Risk and Rationality.* Dordrecht: Kluwer Academic Publishers, 1992, pp. 325–58.

Camerer, Colin F., and Daniel Lovallo, "Optimism and Reference-Group Neglect in Experiments on Business Entry." California Institute of Technology Working Paper No. 975, 1996.

Camerer, Colin F., and Richard Thaler, "Anomalies: Ultimatums, Dictators and Manners," *Journal Economic Perspectives,* Spring 1995, *9*:2, 209–19.

Camerer, Colin F., Eric Johnson, Talia Rymon, and Sankar Sen, "Cognition and Framing in Sequential Bargaining for Gains and Losses." In Binmore, K., A. Kirman, and P. Tani, eds., *Contributions to Game Theory.* Cambridge: Massachusetts Institute of Technology Press, 1993, pp. 27–47.

Camerer, Colin F., Marc J. Knez, and Roberto Weber, "Timing and Virtual Observability in Ultimatum Bargaining and 'Weak Link' Coordination Games." California Institute of Technology Working Paper No. 970, 1996.

Cooper, Russell, Douglas DeJong, Robert Forsythe, and Thomas Ross, "Forward Induction in the Battle-of-the-Sexes Games," *American Economic Review,* December 1993, *83,* 1303–16.

Crawford, Vincent P., "Adaptive Dynamics in Coordination Games," *Econometrica,* January 1995, *63,* 103–43.

Crawford, Vincent P., "Theory and Experiment in the Analysis of Strategic Interaction." In Kreps, David M., and Kenneth F. Wallis, eds., *Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress.* Vol. 1, Cambridge: Cambridge University Press, 1997, pp. 206–42.

Crawford, Vincent P., and Hans Haller, "Learning How to Cooperate: Optimal Play in Repeated Coordination Games," *Econometrica,* May 1990, *58,* 571–95.

Edgeworth, Francis Ysidro, *Mathematical Psychics.* 1881. Reprint, New York: Augustus M. Kelley, Publishers, 1967.

Fehr, Ernst, Georg Kirchsteiger, and Arno Riedl, "Does Fairness Prevent Market Clearing? An Experimental Investigation," *Quarterly Journal of Economics,* May 1993, *108,* 437–59.

Ho, Teck, Colin Camerer, and Keith Weigelt, "Iterated Dominance and Learning in Experimental 'P-Beauty Contest' Games," *American Economic Review,* forthcoming.

Holt, Debra J., "An Empirical Model of Strategic Choice with an Application to Coordination Games," working paper, Queen's University, Department of Economics, 1993.

Kahneman, Daniel, "Experimental Economics: A Psychological Perspective." In Tietz, R., W. Albers, and R. Selten, eds., *Bounded Rational Behavior in Experimental Games and Markets.* Berlin: Springer-Verlag, 1988, pp. 11–18.

Keynes, John Maynard, *The General Theory of Interest, Employment, and Money.* London: Macmillan, 1936.

Kreps, David M., "Corporate Culture and Economic Theory." In Alt, J., and K. Shepsle, eds., *Perspectives on Positive Political Economy.* Cambridge: Cambridge University Press, 1990, pp. 90–143.

Ledyard, John, "Public Goods Experiments." In Kagel, J., and A. Roth, eds., *Handbook of Experimental Economics.* Princeton: Princeton University Press, 1995, pp. 111–94.

McDaniel, Tanga, E. Elisabet Rutström, and Melonie Williams, "Incorporating Fairness into Game Theory and Economics: An Experimental Test with Incentive Compatible Belief Elicitation," working paper, University of South Carolina, Department of Economics, March 1994.

McKelvey, Richard, and Thomas Palfrey, "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior,* July 1995, *10,* 6–38.

Mehta, Judith, Chris Starmer, and Robert Sugden, "The Nature of Salience: An Experimental Investigation of Pure Coordination Games," *American Economic Review,* June 1994, *84,* 658–73.

Nagel, Rosemarie, "Unraveling in Guessing Games: An Experimental Study," *American Economic Review,* December 1995, *85,* 1313–26.

Neale, Margaret A., and Max H. Bazerman, "The Effects of Framing and Negotiator Overconfidence on Bargaining Behaviors and Outcomes," *Academy of Management Journal,* March 1985, *28,* 34–49.

Neale, Margaret A., and Max H. Bazerman, *Cognition and Rationality in Negotiation.* New York: Free Press, 1991.

Preston, C. E., and S. Harris, "Psychology of Drivers in Traffic Accidents," *Journal of Applied Psychology,* 1965, *49:*4, 284–88.

Rabin, Matthew, "Incorporating Fairness into Game Theory and Economics," *American Economic Review,* December 1993, *83,* 1281–302.

Rapoport, Amnon, Darryle A. Seale, Ido Erev, and James A. Sundali, "Equilibrium Play in Large Group Market Entry Games," *Management Science,* forthcoming.

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

Schelling, Thomas, *The Strategy of Conflict.* Cambridge, Mass.: Harvard University Press, 1960.

Sally, David, "Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments from 1958 to 1992," *Rationality and Society,* January 1994, *7,* 58–92.

Stahl, Dale, and Paul Wilson, "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior,* July 1995, *10,* 218–54.

Svenson, Ola, "Are We all Less Risky and More Skillful than our Fellow Drivers?," *Acta Psychologica,* February 1981, *47:*2, 143–48.

Tversky, Amos, and Daniel Kahneman, "Advances in Prospect Theory: Cumulative Representations of Uncertainty," *Journal of Risk and Uncertainty,* October 1992, *5,* 297–323.

Van Huyck, John B., Ray B. Battalio, and Richard O. Beil, "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure," *American Economic Review,* March 1990, *80,* 234–48.

von Neumann, John, and Oskar Morgenstern, *The Theory of Games and Economic Behavior.* Princeton: Princeton University Press, 1944.