minants, as well as the rational, if game theory is to become as descriptively appealing as it is normatively.

Experience and decisions

Edmund Fantino and Stephanie Stolarz-Fantino

Department of Psychology, University of California–San Diego, La Jolla, CA 92093-0109. **efantino@ucsd.edu** **sfantino@psy.ucsd.edu**

Abstract: Game-theoretic rationality is not generally observed in human behavior. One important reason is that subjects do not perceive the tasks in the same way as the experimenters do. Moreover, the rich history of cooperation that participants bring into the laboratory affects the decisions they make.

Colman reviews many instances of game playing in which human players behave much more cooperatively and receive larger payoffs than permitted by conceptions of strict rationality. Specifically, he points out that although "Game-theoretic rationality requires rational players to defect in one-shot social dilemmas" (sect. 6.11), experimental evidence shows widespread cooperation. We agree that strict rationality does not accurately portray or predict human behavior in interactive decision-making situations. Particularly problematic are predictions made on the basis of backward induction. The Chain-store and Centipede games are good examples. In each case, backward induction makes it appear that the likely last move is inevitable, rather than one of a number of possible outcomes, as it must appear to the participant. In any case, it is unlikely that participants would reason backwards from the conclusion, even if such reasoning made sense. For example, Stolarz-Fantino et al. (2003) found that students were more likely to demonstrate the conjunction effect (in which the conjunction of two statements is judged more likely than at least one of the component statements) when the conjunction was judged before the components, than when it was judged after them. Further, if people easily reasoned backward from likely end-states, they should be more adept at demonstrating self-control (preferring a larger, delayed reward to a smaller, more immediate reward) than in fact they are (see discussion in Logue 1988).

Colman proposes "Psychological game theory" as a general approach that can be argued to account for these deviations. We agree that this is a promising approach, although it is a fairly broad and nonspecific approach as presented in the target article. We would add a component to Psychological game theory that appears to be relevant to the types of problems discussed: the pre-experimental behavioral history of the game participants. We are studying various types of irrational and nonoptimal behavior in the laboratory (e.g., Case et al. 1999; Fantino 1998a; 1998b; Fantino & Stolarz-Fantino 2002a; Goodie & Fantino 1995; 1996; 1999; Stolarz-Fantino et al. 1996; 2003) and are finding a pronounced effect of past history on decision-making (a conclusion also supported by Goltz' research on the sunk-cost effect, e.g., Goltz 1993; 1999). One example will suffice.

A case of illogical decision-making is base-rate neglect, first developed by Kahneman and Tversky (1973) and discussed often in this journal (e.g., Koehler 1996). Base-rate neglect refers to a robust phenomenon in which people ignore or undervalue background information in favor of case-specific information. Although many studies have reported such neglect, most have used a single "paper-and-pencil" question with no special care taken to insure attentive and motivated subjects. Goodie and Fantino wondered if base-rate neglect would occur in a behavioral task in which subjects were motivated and in which they were exposed to repeated trials. We employed a matching-to-sample procedure (MTS), which allowed us to mimic the base-rate problem quite precisely (Goodie & Fantino 1995; 1996; 1999; Stolarz-Fantino & Fantino 1990). The sample in the MTS task was either a blue or green light. After sample termination, two comparison stimuli appeared: these were always a blue and a green light. Subjects were instructed to choose either. We could present subjects with repeated trials rapidly (from 150 to 400 trials in less than a one-hour session, depending on the experiment) and could readily manipulate the probability of reinforcement for selecting either color after a blue sample and after a green sample. Consider the following condition (from Goodie & Fantino 1995): Following either a blue sample or a green sample, selection of the blue comparison stimulus is rewarded on 67% of trials, and selection of the green comparison stimulus is rewarded on 33% of trials; thus, in this situation the sample has no informative or predictive function. If participants responded optimally, they should have come to always select blue, regardless of the color of the sample; instead they focused on sample accuracy. Thus, after a green sample, instead of always choosing blue (for reward on 67% of trials) they chose the (matching) green comparison stimulus on 56% of trials (for a 48% rate of reward). This continued for several hundred trials. In contrast, Hartl and Fantino (1996) found that pigeons performed optimally, ignoring the sample stimulus when it served no predictive function. They did not neglect base-rate information.

What accounts for pigeons' and people's differing responses to this simple task? We have speculated that people have acquired strategies for dealing with matching problems that are misapplied in our MTS problem (e.g., Stolarz-Fantino & Fantino 1995). For example, from early childhood, we learn to match like shapes and colors at home, in school, and at play (e.g., in picture books and in playing with blocks and puzzles). Perhaps, this learned tendency to match accounts for base-rate neglect in our MTS procedure. If so, Goodie and Fantino (1996) reasoned that base-rate neglect would be eliminated by using sample and comparison stimuli unrelated to one another (line orientation and color). In this case, base-rate neglect was indeed eliminated. To further assess the learning hypothesis, Goodie and Fantino (1996) next introduced an MTS task in which the sample and comparison stimuli were physically different but related by an extensive history. The samples were the words "blue" and "green"; the comparison stimuli were the colors blue and green. A robust base-rate neglect was reinstated. Ongoing research in our laboratory is showing that pigeons with sufficient matching experience (where matching is required for reward) can be induced to commit base-rate neglect. These and other studies have led us to conclude that base-rate neglect results from preexisting learned associations.

How might learned associations account for nonoptimal decisions in the Prisoner's Dilemma Game (PDG)? Rationality theory argues that the selfish response is optimal. But we have been taught since childhood to be unselfish and cooperative. For many of us, these behaviors have been rewarded with praise throughout our lives (see the discussion of altruism in Fantino & Stolarz-Fantino 2002b; Rachlin 2002). Moreover, actual deeds of unselfish and cooperative behavior are often reciprocated. Why then should these behaviors not "intrude" on the decisions subjects make in the laboratory? Viewed from this perspective, there is nothing surprising about the kinds of behavior displayed in PDG. Indeed, such behavior is variable (many subjects cooperate, many defect), as one would expect from the variable behavioral histories of the participants.

A critique of team and Stackelberg reasoning

Herbert Gintis

Emeritus Professor of Economics, University of Massachusetts, Northampton, MA 01060; External Faculty, Santa Fe Institute, Santa Fe, NM.
hgintis@comcast.net **http://www-unix.oit.umass.edu/~gintis**

Abstract: Colman's critique of classical game theory is correct, but it is well known. Colman's proposed mechanisms are not plausible. Insufficient reason does what "team reasoning" is supposed to handle, and it applies to a broader set of coordination games. There is little evidence ruling out more traditional alternatives to Stackelberg reasoning, and the latter is implausible when applied to coordination games in general.

Colman's critique of classical game theory is correct, but it is well known. He misses one critique that I consider to be among the most telling. If two "rational" players play a game with a unique, strictly mixed strategy equilibrium, neither player has an incentive to play using this equilibrium strategy, because in a true one-shot game, there is absolutely no reason to randomize. It is easy to explain why one would prefer that one's opponent not know which action we will take, and it is possible to work this up into a full-fledged justification of randomizing. But in a true one-shot, your opponent knows nothing about you, so even if you choose a pure strategy, you do no worse than by randomizing. The evolutionary game-theoretic justification is that in a large population of agents meeting randomly and playing the game in each period, in equilibrium a fraction of the population will play each of the pure strategies in proportion to that strategy's weight in the mixed-strategy Nash equilibrium.

Indeed, most of the problems with classical game theory can be handled by evolutionary/behavioral game theory, and do not need models of "nonstandard reasoning" (Gintis 2000). For instance, in a pure coordination game with a positive payoff-dominant equilibrium, and the payoffs to noncoordinated choices zero, evolutionary game theory shows that each pair of coordinated choices is a stable equilibrium, but if there are "trembles," then the system will spend most of its time in the neighborhood of the payoff-dominant equilibrium (Young 1993).

As Colman notes, many of the empirical results appearing to contradict classical game theory, in fact contradict the assumption that agents are self-regarding. In fact, agents in many experimental situations care about fairness, and have a propensity to cooperate when others cooperate, and to punish noncooperators at personal cost, even when there can be no long-run personal material payoff to so doing. For an analysis and review of the post-1995 studies supporting this assertion, see Gintis 2003.

Evolutionary game theory cannot repair all the problems of classical game theory, because evolutionary game theory only applies when a large population engages in a particular strategic setting for many periods, where agents are reassigned partners in each period. We still need a theory of isolated encounters among "rational" agents (i.e., agents who maximize an objective function subject to constraints). Colman proposes two such mechanisms: team reasoning and Stackelberg reasoning. I am not convinced that either is a useful addition to the game-theoretic repertoire.

Concerning "team reasoning," there is certainly much evidence that pregame communication, face-to-face interaction, and framing effects that increase social solidarity among players do increase prosocial behavior and raise average group payoffs, but this is usually attributed to players' placing positive weight on the return to others, and increasing their confidence that others will also play prosocially. But these are nonstandard *preference* effects, not nonstandard *reasoning* effects. Choosing the payoff-maximum strategy in pure coordination games, where players receive some constant nonpositive payoff when coordination fails, is most parsimoniously explained as follows. If I know nothing about the other players, then all of my strategies have an equal chance of winning, so personal payoff maximization suggests choosing the payoff maximum strategy. Nothing so exotic as "team reasoning" is needed to obtain this result. Note that if a player *does* have information concerning how the other players might choose, an alternative to the payoff-maximum strategy may be a best response.

Moreover, "team reasoning" completely fails if the pure coordination game has nonconstant payoffs when coordination is not achieved. Consider, for instance, the following two-person game. Each person chooses a whole number between 1 and 10. If the numbers agree, they each win that amount of dollars. If the numbers do not agree, they each lose the larger of the two choices. For example, if one player chooses 10, and the other chooses 8, they both lose ten dollars. This is a pure coordination game, and "team reasoning" would lead to both players choosing 10. However, all pure strategies are evolutionary equilibria, and computer simulation shows that the higher numbers are less likely to emerge when the simulation is randomly seeded at the start (I'll send interested readers the simulation program). Moreover, if an agent knows nothing about his partner, it is easy to show, using the Principle of Insufficient Reason, that 2 and 3 have the (equal and) highest payoffs. So if an agent believes that partners use the same reasoning, he will be indifferent between 2 and 3. By the same reasoning, if one's partner chooses 2 and 3 with equal probability, then the payoff to 3 is higher than the payoff to 2. So 2 is the "rational" choice of "ignorant" but "rational" agents.

Colman argues that there is strong evidence supporting Stackelberg reasoning, but he does not present this evidence. Some is unpublished, but I did look at the main published article to which he refers (Colman & Stirk 1998). This article shows that in 2 × 2 games, experimental subjects overwhelmingly choose Stackelberg solutions when they exist. However, a glance at Figure 1 (p. 284) of this article shows that, of the nine games with Stackelberg solutions, six are also dominance-solvable, and in the other three, any reasoning that would lead to choosing the payoff-maximum strategy (including the argument from insufficient reason that I presented above), gives the same result as Stackelberg reasoning. So this evidence does not even weakly support the existence of Stackelberg reasoning. I encourage Colman to do more serious testing of this hypothesis.

I find the Stackelberg reasoning hypothesis implausible, because if players used this reasoning in pure coordination games, it is not clear why they would not do so in other coordination games, such as Battle of the Sexes (in this game, both agents prefer to use the same strategy, but one player does better when both use strategy 1, and the other does better when both use strategy 2). Stackelberg reasoning in this game would lead the players never to coordinate, but always to choose their preferred strategies. I know of no experimental results using such games, but I doubt that this outcome would be even approximated.

How to play if you must

Hans Haller

Department of Economics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061. **haller@vt.edu**

Abstract: Beyond what Colman is suggesting, some residual indeterminacy of Nash equilibrium may remain even after individual rationality is amended. Although alternative solution concepts can expand the positive scope (explanatory power) of game theory, they tend to reduce its accuracy of predictions (predictive power). Moreover, the appeal of alternative solutions may be context-specific, as illustrated by the Stackelberg solution.

Analysis of a strategic or noncooperative game presumes that the players are committed to participate. Normative analysis then aims at an unambiguous recommendation of how to play the game. If the analyst and the players adhere to the same principles of rationality, then the players will follow the recommendation; indeed, the players can figure out how to play without outside help. But can they? Like Colman, I shall refrain from elaborating on bounded rationality.

Andrew Colman presents the argument of Gilbert, that common knowledge of individual rationality does not justify the use of salient (exogenous, extrinsic) focal points to resolve indeterminacy. Nor does it justify endogenous or intrinsic focal points based on payoff dominance or asymmetry. This argument is in line with the critique by Goyal and Janssen (1996) of Crawford and Haller's heuristic principle, to stay coordinated once coordination is obtained. It applies as well to folk theorem scenarios, as in the infinitely repeated Prisoner's Dilemma Game (PDG): None of the multiple equilibria is distinguished on the grounds of individual rationality alone. The argument shows that principles other than individual rationality have to be invoked for equilibrium selection