

Game theory need not abandon individual maximization

John Monterosso^a and George Ainslie^b

^aDepartment of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, CA 90024;

^bDepartment of Psychiatry, Coatesville VA Medical Center, Coatesville, PA 19320. jmont@ucla.edu george.ainslie@med.va.gov

Abstract: Colman proposes that the domain of interpersonal choice requires an alternative and nonindividualistic conception of rationality. However, the anomalies he catalogues can be accounted for with less radical departures from orthodox rational choice theory. In particular, we emphasize the need for descriptive and prescriptive rationality to incorporate recursive interplay between one's own choices and one's expectation regarding others' choices.

Colman proposes that an alternative conception of rationality is required to account for human interaction, and he provides some suggestions in this direction. What he specifically sees the need to give up is "methodological individualism" – the premise that "rational play in games can be deduced, in principle, from one-person rationality considerations" (Binmore 1994a, quoted in target article, sect. 4.1, para. 1). We think the anomalies he catalogues can be accounted for without abandoning this foundational principle of deterministic behavioral science.

First, the prevailing payoffs in experimental games are not the same as the specified payoffs. Social interactions are rife with invisible contingencies that are impossible to bring under full experimental control. Human beings are fundamentally social creatures, which entails the presence of powerful interpersonal motivations, too numerous to list. Otherwise, anomalous play, such as rejecting a low offer in a one-shot ultimatum game or cooperating in a one-round prisoner's dilemma game, is sensible if we allow that the dollars offered do not exhaust the prevailing payoffs. Colman discusses this type of proposal (Camerer's "behavioral game theory," Rabin's fairness equilibrium), but he concludes it is not enough to account for all the phenomena he presents. We agree, and furthermore, do not think that the subjects' many motives, beyond maximizing the specified matrix outcomes, are orderly enough to inspire any useful addition to game theory (such as adding X points to particular cells); discrepancies between the specified payoffs and the prevailing payoffs will always be noise in the experiment, the friction that distorts the ideal physics lab.

However, permitting the free use of probability estimates of other's choices should be enough to let "methodological individualism" both describe and prescribe rationality to the extent that subjects *are* motivated by the specified matrices of the game. Of particular importance in explaining otherwise anomalous play is the subject's use of her own inclinations and behavior as test cases that inform her expectations regarding what others will do. In the kinds of situations game theorists care about, it is neither descriptively tenable *nor prescriptively effective* to require individuals to finalize their assessments of what others will do prior to considering what they will do. Instead, we think that a rational individual is simultaneously engaging in both computing expectation of what the other player will be motivated to do and contemplating what she herself should do, and each process informs the other. In the absence of specific information about one's counterpart, what better basis is there to predict her behavior than via one's own response to the situation?

Colman describes such a recursive process in characterizing one attempt to develop a game-theoretic rationale for the Pareto-dominant H-H solution in the Hi-Lo game. In this account, Player I assumes by default that Player II's strategies are equally probable. She thus concludes she should choose H, because the probability-weighted sum is higher. But this, Colman adds, violates rational choice theory. "By the transparency of reason, Player I's intention to choose H would be common knowledge and would induce Player II to choose the best reply, namely H, with *certainty*, contradicting Player I's initial assumption" [i.e., of equal probabil-

ity of moves] (sect. 5.6, para. 4). While such recursion may violate game theory's constraints, we think it is descriptively accurate, prescriptively rational, and it does not entail abandoning methodological individualism.

The recursion between someone's own perceived choices and their expectations about the choices of others is easier to see in a *discoordination* variant of the Hi-Lo game, in which players get to keep their choice (in some monetary unit) if and only if they chose differently from each other. With no a priori expectation regarding what Player II will choose, Player I's first-order inclination is to choose the high amount, following the same logic as above. But seeing the similarity of her counterpart's predicament, she may expect her to have thought the same way, giving her the second-order inclination that she must go for the lower amount to get anything. But then again, if she thinks her counterpart is a similarly sophisticated sort, she might get the feeling that her counterpart went through the same thought process, thus giving her the third-order inclination that maybe she *should* therefore go for H. The more similar she thinks her counterpart to be to herself, the more dizzying the potential for iteration, and the less likely there will be a probable solution.

The recursive prediction model has the advantage that it also provides intertemporal bargaining within the individual person. In situations that involve resisting temptation, individuals cannot be certain of their own future choices. The need to choose in the present, with an eye to the precedent this choice will set for the future (e.g., whether or not I am sticking to my diet), places people in a situation analogous to a repeated prisoner's dilemma (PD) game, as we have argued elsewhere (Ainslie 2001, pp. 90–104; Ainslie & Monterosso 2003; Monterosso et al. 2002). Briefly, the danger that future selves will see past violations of a resolution as a reason to violate it, in turn, is similar to the danger that one player's defection will cause the other(s) to defect. But, in this bargaining, a person may propose a choice to herself ("I'll have an ice cream"), then put herself in the shoes of her future self to evaluate it retrospectively ("I'll have gone off my diet"), then revise her current choice in light of this evaluation ("I'll have a muffin instead"), and evaluate this ("no"), and propose again ("a bran muffin") at some length before making a single concrete choice. Choices may turn out to divide along salient features, just as in the Hi-Lo game, not because of their intrinsic payoff, but because they make intertemporal cooperation more likely. Intertemporal bargaining theory predicts the emergence of both positive and negative features that have been ascribed to willpower. It generates an internal version of Adam Smith's "unseen hand" without assuming an innate faculty of self-control.

Second-order indeterminacy

Marco Perugini

Department of Psychology, University of Essex, Colchester, CO4 3SQ United Kingdom. mperug@essex.ac.uk
<http://privatewww.essex.ac.uk/~mperug>

Abstract: Psychological game theory, as defined by Colman, is meant to offer a series of solution concepts that should reduce the indeterminacy of orthodox game theory when applied to a series of situations. My main criticism is that, actually, they introduce a second-order indeterminacy problem rather than offering a viable solution. The reason is that the proposed solution concepts are under-specified in their definition and in their scope.

Colman looks at game theory from a psychological perspective. In the first part of his article, he convincingly argues about the limitations of orthodox game theory, especially when applied to social interactions. The examples are well chosen and the case is well built. This is an important contribution that might help us to focus, once and for all, on these important issues. However, Colman's suggestion of psychological game theory as a way forward to overcome the severe limitations of orthodox game theory in ex-