nitive mechanisms produce decisions, which are sometimes "individually rational," sometimes "collectively rational," and sometimes "not rational at all." Because these mechanisms have been evolved and developed to assure human survival, they will, most of the time, produce results that are "rational" or "optimal" from some point of view – this is what makes rationality a good naive theory. However, this does not mean that people explicitly follow the rules of maximization prescribed by the theory.

Colman proposes an eclectic collection of ad-hoc strategies (team reasoning, Stackelberg reasoning, epistemic, and nonmonotonic reasoning), which are all different forms of explicit deductive reasoning. Deduction can certainly play a role in decision-making, but it is not enough to explain it. Recent studies revealed that analogy-making is a more basic mechanism of human thinking, which is present from early infancy and is used ubiquitously in everyday life (Gentner et al. 2001). Analogy-making is a process of perceiving one situation (target) in terms of another (base), thereby preserving the system of relations among elements and transferring knowledge from the base to the target. Arguments have been presented that deduction is in fact based on analogy, and a special form of it (Halford 1993; Kokinov 1992). Markman and Moreau (2001) have reviewed the evidence that analogy plays an important role in perceiving and framing the decision situation, as well as in comparison of the alternatives. Moreover, analogy may be used both explicitly and implicitly (Kokinov & Petrov 2001; Markman & Moreau 2001). Thus, analogy may play a unifying role in describing the mechanisms of decision-making.

Analogy-making may explain the paradoxes of using the focal points described by Colman. They are easily perceivable and analogous to focal points in other games. Therefore, it is natural to expect people to use them again and again if previous experience of using a focal point has been successful. Similar arguments may be applied to social dilemmas and trust games. If another player has used a certain strategy in a previous case, I may expect him or her to behave the same way in an analogous situation, and thus have a prediction for his or her behavior.

Analogies may be applied at various levels: Analogies to previous cases of decision-making in the same game or analogies to games with similar structure; analogies to cases of social interaction with the same individual or to cases of social interactions with individuals who are considered analogous (i.e., are in similar relations to me, like family or team members). Thus, even a novice in a particular game can still use his or her previous experience with other games.

Analogy can explain the "deviations" from the prescribed "rational" behavior and the individual differences among players. If a player has an extensive positive experience of cooperative behavior (i.e., many successful cases of benefiting from acting together), and if the current game is found to be analogous to one of these cases, then he or she might be expected to act cooperatively (even if this is not the optimal strategy). On the contrary, if the game reminds the player of a previous case of betrayal or fraud, then defection strategy should be expected.

In summary, analogy may play a crucial role in a future theory of decision-making. Instead of explaining rationality with rules for utility maximization, which people follow or break, we may explain human behavior by assuming that decisions are made by analogy with previous cases (avoid strategies that were unsuccessful in analogous situations and re-use strategies that were successful). Thus, utility maximization is an emergent property that will emerge in most cases, but not always. In this view, rationality is an emergent phenomenon, and rational rules are only a rough and approximate explanation of human behavior.

## Wanted: A reconciliation of rationality with determinism

Joachim I. Krueger

*Department of Psychology, Brown University, Providence, RI 02912.*
**joachim_krueger@brown.edu**
**http://www.brown.edu/departments/psychology/faculty/krueger.html**

**Abstract:** In social dilemmas, expectations of reciprocity can lead to fully determined cooperation concurrent with the illusion of choice. The choice of the dominant alternative (i.e., defection) may be construed as being free and rational, but only at the cost of being incompatible with a behavioral science claiming to be deterministic.

The conspicuous failure of orthodox game theory is its inability to account for cooperative behavior in noniterated social dilemmas. Colman outlines a psychological revision of game theory to enhance the predictability of hitherto anomalous behavior. He presents the Stackelberg heuristic as a form of evidential reasoning. As Colman notes, evidential reasoning is assumed to lead respondents to shun the dominating alternative in Newcomb's problem and in decisions to vote. In the prisoner's dilemma game (PDG), however, Stackelberg reasoning leads to defection (Colman & Stirk 1998). Thus, Stackelberg reasoning appears to be neither evidential nor parsimonious in this domain. After all, players can select the dominating alternative in the PDG without making any predictions of what their opponents will do. How, then, can evidential reasoning lead to cooperation?

The logic of the PDG is the same as the logic of Newcomb's problem (Lewis 1979). Just as players may expect that their choices will have been predicted by Newcomb's savvy demon, they may expect that their choices in the PDG will most likely be matched by their opponent's choices (unless the rate of cooperation is exactly 50%). The issue is whether this statistical realization gives cooperators (or one-boxers, in Newcomb's case) license to lay claim to being rational.

Orthodox game theorists insist on defection, because a player's cooperation cannot make an opponent's cooperation more likely. Evidentialists, however, claim that cooperation may be chosen without assuming a causal effect on the opponent's choice. Only the assumption of conditional dependence is needed. If nothing is known about the opponent's choice, conditional dependence is obvious *after* a player committed to a choice. By definition, most players choose the more probable alternative, which means that the choices of two independent players are more likely to be the same than different (Krueger 1998). Because time is irrelevant, it follows that it is more likely that two players *will* make the same, instead of different, choices. In the extreme case, that players expect their responses to be reciprocated without fail, their dilemma devolves into a choice between mutual cooperation and mutual defection. As mutual cooperation offers the higher payoff, they may choose cooperation out of self-interest alone.

Evidentialist reasoning is distasteful to the orthodox mind because it generates two divergent conditional probabilities that cannot both be correct (i.e., $p$[opponent cooperation/own cooperation] $>$ $p$[opponent cooperation/own defection]). Choosing the behavior that is associated with the more favorable prospect then smacks of magical thinking. But causal assumptions enter at two levels: at the level of the investigator and at the level of the participant. Investigators can safely assume that players' efforts to influence their opponents are pointless. Players, however, may *think* they can exert such influence. Although this expectation is irrational, it does not invalidate their cooperative choices. Note that investigators can also subscribe to a more plausible causal argument, which holds that both players' choices result from the same set of latent variables. These variables, whatever they may be, produce the proportions of cooperation found in empirical studies. Players who realize that one option is more popular than the other, but do not know which, can *discover* the popular choice by observing their own. The fact that they may have an experience of


https://doi.org/10.1017/S0140525X03380057 Published online by Cambridge University Press

unfettered choice, and perhaps even hope to influence their opponents, is quite irrelevant (Wegner 2002).

The burgeoning literature on social dilemmas suggests that individual behavior in these situations presents a more poignant dilemma to the investigators than to the participants. However modest their predictive successes may be, experimental studies of social behavior rest on a bedrock assumption of determinism. In this spirit, experimentalists assume that individuals' judgments and decisions are fully determined (Bargh & Ferguson 2000). It is ironic that research participants who are cast into the PDG or confronted with Newcomb's problem can satisfy norms of rationality only by denying any determining effect on their own behavior that would make them act like most others.[1] They are enjoined to choose defection without drawing any inference as to what this might say about their opponents' choices. Evidentialists, in contrast, can maintain a deterministic outlook without being perplexed. They need only assume that cooperators choose "as if" they were free.

Incidentally, players working on the assumption that their own choices will likely be reciprocated are also comfortable with common-interest games. They do well without experiencing the puzzlement of orthodox game theorists and even without resorting to von Stackelberg's best-bet heuristic. Perhaps more importantly, evidential reasoning preserves *methodological individualism* in common-interest games. Collective preferences, as entailed by team spirit, are unnecessary. A game in which players are paid only if their choices do not match, however, would be a true puzzle to the evidentialist and the orthodox alike. Even team spirit, no matter how lofty its intent, cannot overcome this hurdle.

NOTE

**1.** In iterated PDGs, the assumption of determinism is more apparent than in one-shot games. Players' choices are assumed to be controlled by the design of the game (i.e., the experimenters) and by each other's choices in preceding rounds (e.g., Rachlin 2002).

## Let's cooperate to understand cooperation

John Lazarus

*Evolution and Behaviour Research Group, Psychology, School of Biology, University of Newcastle, Newcastle upon Tyne, NE2 4HH United Kingdom.*
**j.lazarus@ncl.ac.uk**
**http://www.ncl.ac.uk/biol/staff/john_lazarus.html**

**Abstract:** The importance of understanding human cooperation urges further integration between the relevant disciplines. I suggest ideas for bottom-up and top-down integration. Evolutionary psychology can investigate the kinds of reasoning it was adaptive for humans to employ. Disciplines can learn from each other's approaches to similar problems, and I give an example for economics and evolutionary biology.

Understanding the factors that facilitate and constrain human cooperation is of the greatest importance. I suggest here ways in which disciplines with a convergent interest in cooperation might fruitfully interact, with an emphasis on theoretical modelling.

Colman describes "nonstandard forms of reasoning" that help to explain irrational social decisions. Psychological game theory should employ the methods of evolutionary psychology (Tooby & Cosmides 1992) to determine both the kinds of social problems that early humans were selected to solve, and the kinds of reasoning that were adaptive to employ. Such an analysis of social problems has shown that human reasoning is well-designed for cheater detection, for example (Cosmides & Tooby 1992). An evolutionary analysis of kinds of reasoning could start with team reasoning (target article, sect. 8.1), for which two potential adaptive explanations seem worth pursuing. Team reasoning might be favoured where cooperation benefits the group, or where maximizing collective payoff raises one's reputation and thus brings future rewards (Milinski et al. 2002). Evolutionary game theory is the tool for analyzing the evolutionary fate of competing modes of reasoning.

Knowledge of social decision-making in dyads and small, unstructured groups is a starting point for understanding cooperation at the higher levels of structured groups, firms, institutions, communities, and states (cf. Hinde 1987). Table 1 (see overleaf) lists disciplines sharing an interest in cooperation, indicating their interests, methods, and levels of analysis; it is not exhaustive (e.g., nothing on military strategy). Its purpose is to indicate the multidisciplinary nature of cooperation, to encourage further interdisciplinary work (following, e.g., Axelrod 1984; 1997; Frank 1988), and to act as a reference point for the following proposals in this direction.

Colman shows that there is much to be done before we understand cooperative decision-making at the lowest level, although understanding should be advanced by reference to the social psychological foci in Table 1. To bring greater psychological reality to decision theory in the structured groups of institutions and societies, game theory models and psychological game theory findings should be combined with the decision-making models of economics and related disciplines (Table 1; see also Axelrod 1997).

This bottom-up approach should be complemented by psychological game theory adopting top-down insights gained from analyses of real-life economic behaviour. Decision-making in these real-life contexts may reflect evolved predispositions, and may tap motivations at work even in the economically elementary scenarios of the psychological laboratory. For example, studies of the way in which communities govern their own use of common pool resources (CPRs), such as grazing pastures (Ostrom 1990), may reveal evolved influences on cooperative decision-making, and even evolved modes of reasoning, because the hunting and gathering activities of early humans also have CPR properties. Successful CPR decisions are characterized by: a clear in-group/out-group distinction; resource provision in proportion to need and sharing of costs in proportion to ability to pay; and graded punishments for the greedy (Ostrom 1990). Whether these characteristics apply to decision-making in other kinds of cooperative relationship is open to evolutionary psychological and empirical analysis. It would be valuable to know whether cooperation was rational and evolutionarily stable in CPR scenarios.

In addition to bottom-up and top-down integration, different disciplines can surely learn from each other's approaches to similar problems. I close with an example. In economics, a common pool resource is "subtractable," because resources removed by one person are unavailable for others. In contrast, a pure public good (e.g., a weather forecasting system) is "nonsubtractive" in that its use by one person leaves it undiminished for others (Ostrom 1990, pp. 31–32). In evolutionary biology, parental investment in offspring is of two kinds, "shared" and "unshared," respectively, the identical concepts just described from economics. Food for the young must be shared among them, whereas parental vigilance for predators is enjoyed by all simultaneously. Modelling in the evolutionary biology case has examined the influence of the number of users on the optimal allocation of investment, and on conflict between producer (parent) and user (offspring) (Lazarus & Inglis 1986). Could economists use these results? Have economists produced similar results that evolutionary biologists should know about?

ACKNOWLEDGMENTS

I am grateful to the "Society and Modernization" seminar group at the University of Newcastle for broadening my horizons.