Rusbult & Van Lange 2003). This taxonomy allows one to characterize interaction situations. For example, a social dilemma would be characterized as one involving relatively high levels of interdependence, based on unilateral actions of the partner, and also characterized by a fairly strong conflict of interest; and social dilemmas may differ in terms of temporal structure (e.g., single-trial vs. iterated) and information availability (e.g., complete or incomplete information regarding one another's preferences).

Second, interdependence theory assumes that the outcomes in any interaction situation ("the given preferences") may be psychologically *transformed* into a subjective situation representing effective preferences, which are assumed to guide behavior and ultimately social interaction. Examples of transformation rules are interaction goals such as enhancement of both one's own and other's outcomes (MaxJoint), equality in outcomes (MinDiff), other's outcomes (MaxOther), or relative advantage over other's outcomes (MaxRel). Although transformations may be a product of careful reasoning and thought, they may also occur in a fairly automatic manner, involving very little thought or deliberation. As such, transformations deviate not only from self-interest, but also from rationality, in that individuals are not assumed to obey criteria of strict rationality. More importantly, transformations are assumed to accompany cognitive and affective processes in guiding behavior and shaping interaction (see Kelley et al. 2003; Rusbult & Van Lange 2003).

Finally, interdependence theory focuses on both individual and collective levels of analyses, in that it explicitly seeks to understand *social interaction,* which is conceptualized as a product of two individuals (with their basic preferences and transformational tendencies) and the interaction situation. Social interactions are also assumed to shape relatively stable embodiments of transformations, at the intrapersonal level (i.e., dispositions such as social value orientation), at the relationship level (i.e., partner-specific orientations, such as commitment), and at the cultural level (i.e., broad rules for conduct, such as social norms; see Rusbult & Van Lange 2003; Van Lange et al. 1997).

Interdependence theory and psychological game theory. As noted earlier, the parts of psychological game theory dealing with preference seem to be well-captured by interdependence theory. Needless to say, the notion of transformation explicates deviations of rationality and self-interest, and it is a theory that focuses on both individual and collective levels of analysis. Moreover, although transformations are individual-level rules, they do have strong implications for the collective level of analysis. For example, transformations may be interrelated with group-based variables, such as group identification ("we-thinking"), group attachment, or feelings of group responsibility. A case in point is the demonstration that individuals with prosocial orientation define rationality at the collective level, not at the individual level, thus judging cooperation as more intelligent than noncooperation in social dilemmas (cf. goal-prescribes-rationality principle, Van Lange 2000). Also, interdependence theory emphasizes the conceptual importance of beliefs, expectations, and interpersonal trust. Following the seminal work of Kelley and Stahelski (1970), the transformations that people actually adopt are assumed to be strongly conditioned by trust, beliefs, and expectations regarding the transformations pursued by particular interaction partners.

Utility of a transformational analysis. We should also briefly comment on Colman's suggestion that a transformational analysis is not especially helpful in understanding the Hi-Lo Matching game. Let us analyze this particular interaction situation for five transformations. First, a transformational analysis indicates that the orientations toward enhancing one's own outcomes (individualism), joint outcomes (cooperation), and a partner's outcomes (altruism) prescribe matching, and more strongly so for Heads than for Tails. Second, mere enhancement of relative advantage (competition) and equality in outcomes (egalitarianism) are irrelevant in this particular situation, because all four cells present equal outcomes for self and other. Given that cooperation and individualism are prevalent orientations, and given that often cooperation is accompanied by egalitarianism (see Van Lange 1999), the transformational analysis indicates that most people will be oriented toward matching Heads (followed by matching Tails). Pure forms of competition or egalitarianism lead to indifference, which in fact may hinder effective coordination between two individuals.

Thus, a transformation analysis may very well account for the fact that people tend to be fairly good at coordinating in the Hi-Lo Matching game and related situations. At the same time, the transformational analysis suggests that one reason people may not be able to coordinate is that at least one individual is merely interested in outperforming the interaction partner (or, less likely, merely interested in obtaining equality in outcomes). More generally, a comprehensive transformation analysis (which includes not only cooperation, as discussed by Colman) helps us understand this specific situation, even though we agree with Colman's implicit assumption that a transformational analysis is typically more strongly relevant to motivational dilemmas, involving more pronounced conflicts among cooperation, equality, individualism, and competition.

Conclusion. Over the past 25 years, interdependence theory has inspired several programs of research in areas as diverse as relationships, norms and roles, interpersonal dispositions, social dilemmas, group decision-making, and negotiation. It is a comprehensive, logical, and psychological theory of social interaction, thereby, to some degree, using the language (and logic) of game theory. Our discussion indicates that the parts of psychological game theory dealing with preference (and illustrated by team reasoning) do not extend interdependence theory in terms of theoretical potential, logic (including parsimony), or psychological breadth. Perhaps the contribution of Colman's article is more strongly rooted in conceptualizing specific lines of reasoning, such as Stackelberg reasoning, and reasoning focusing on common beliefs and nonmonotonic reasoning, which tend to deviate from self-interest or rationality. We are most confident about one broad message that Colman's article shares with interdependence theory – that is, the conviction that "bridging" social psychology with game theory is essential to the further development of the science of social interaction. After all, one needs games ("the situational structure") and the psychology of two individuals ("the processes," i.e., transformations, along with cognition and affect) to understand social interaction.

Toward a cognitive game theory

Ivaylo Vlaev[a] and Nick Chater[b]

[a]*Department of Experimental Psychology, University of Oxford, Oxford, OX1 3UD, United Kingdom;* [b]*Institute for Applied Cognitive Science, Department of Psychology, University of Warwick, Coventry, CV4 7AL, United Kingdom.*
ivaylo.vlaev@psy.ox.ac.uk nick.chater@warwick.ac.uk

Abstract: We argue that solving the heterogeneous problems arising from the standard game theory requires looking both at reasoning heuristics, as in Colman's analysis, and at how people represent games and the quantities that define them.

Colman's elegant and persuasive article describes psychological game theory by introducing formal principles of reasoning, and focuses on several nonstandard reasoning processes (team reasoning, Stackelberg reasoning, and epistemic and nonmonotonic reasoning). The goal is to explain psychological phenomena in game-playing that orthodox game theory, and its conventional extensions, cannot explain. We argue that, in addition, a model is needed of how the economic agent perceives and mentally represents the game initially, before any (strategic) reasoning begins. For instance, the perceived utility of various outcomes might change depending on the previous games seen.

As an illustration of such a possible model, here we offer some initial results from a research program that aims to ground accounts of rationality in general, and decision theory in particular, on the underlying cognitive mechanisms that produce the seemingly paradoxical behavior. Colman's sections 2 and 4 discuss the basic underlying assumptions of expected utility theory and game theory. Existing models of rational choice and interactive game-theoretic decision making typically assume that only the attributes of the game need be considered when reaching a decision; that is, these theories assume that the utility of a risky prospect or strategy is determined by the utility of the outcomes of the game, and transforms the probabilities of each outcome. Decisions are assumed to be based on these utilities.

Our results demonstrate, however, that the attributes of the previously seen prospects and games influence the decisions in the current prospect and game, which suggests that prospects and games are not considered independently of the previously played ones (Stewart et al., in press; Vlaev & Chater, submitted). In particular, Stewart et al. (in press) have revealed the phenomena of "prospect relativity": that the perceived value of a risky prospect (e.g., "p chance of x") is relative to other prospects with which it is presented. This is counter to utility theory, according to which the perceived value of each prospect should be dependent only on its own attributes. Stewart et al. suggest that this phenomenon arises in the representation of the magnitudes that define the prospects, and suggest that the phenomenon has a common origin with related effects in the perception of sensory magnitudes (Garner 1954; Laming 1997; Lockhead 1995).

We have found similar effects, providing a new type of anomaly for orthodox game theory. People play repeated one-shot Prisoner's Dilemma games (Vlaev & Chater, submitted). The degree to which people cooperate in these games is well-predicted by a function of the pay-offs in the game, the cooperation index as proposed in Rapoport & Chammah (1965). Participants were asked on each round of the game to predict the likelihood that their coplayer will cooperate, and then to make a decision as to whether to cooperate or defect. The results demonstrated that the average cooperation rate and the mean predicted cooperation of the coplayer in each game strongly depend on the cooperativeness of the preceding games, and specifically on how far the current game was from the end-points of the range of values of the cooperation index in each session. Thus, in games with identical cooperation indices, people cooperated more and expected more cooperation in the game with higher rank position relative to the other cooperation index values in the sequence. These findings present another challenge to the standard rational choice theory and game theory, as descriptive theories of decision-making under uncertainty, and also to other theories where games are independently considered.

Our proposed account for these results, and also for other problems related to the independence assumption, is that people have poor notions of absolute cooperativeness, risk, and utility, and instead make their judgments and decisions in relative terms, as is described in some existing psychophysical and cognitive theories of perception and judgment of information about magnitudes (intensities of stimulus attributes). Thus, this account departs fundamentally from previous work in this field, by modeling the highly flexible and contextually variable way in which people represent magnitudes, like sums of money, probabilities, time intervals, cooperativeness, and so forth, rather than by assuming that these can be represented on absolute internal psychological scales (i.e., even if these scales exist, they stretch or contract depending on the other stimuli in the environment). We conjecture that the results from the two studies presented here suggest that people use context as a sole determinant of the utility of a strategy, which is a form of a more ecologically adaptive rationality, and therefore any descriptive account of game-theoretic behavior, especially in sequential social dilemmas, should incorporate a model of agents' lower-level cognitive perceptual processes.

This discussion does not answer the paradoxes posed in the target article, but here we would like to make the stronger claim that there are many more phenomena that the standard approach cannot explain (and there will be more to be discovered), and that in order to develop a decent account of human decision behavior in games, a much more radical approach is needed. Our results imply that Colman's version of psychological game theory, as based only on nonstandard forms of reasoning, needs to be supplemented by a more general "cognitive game theory," which grounds decision-making in the underlying cognitive mechanisms that produce the decision behavior. Such a cognitive approach could also include collective rationality criteria (which, as Colman states, are lacking in the standard decision theory), because, for example, categorization of the coplayer as being very similar to me could strongly affect my common belief in each other's rationality, or at least in the similarity of the reasoning processes that we would employ. Also, the perception of the players and the game as being similar to a previous interactive situation, in which the coplayers acted in a certain way (e.g., chose a certain focal point), would enforce the belief that, in the current situation, the coplayers would act in a similar way.

From rationality to coordination

Paul Weirich

Philosophy Department, University of Missouri, Columbia, MO 65211.
weirichp@missouri.edu
http://web.missouri.edu/~philwww/people/weirich.html

Abstract: Game theory's paradoxes stimulate the study of rationality. Sometimes they motivate the revising of standard principles of rationality. Other times they call for revising applications of those principles or introducing supplementary principles of rationality. I maintain that rationality adjusts its demands to circumstances, and in ideal games of coordination it yields a payoff-dominant equilibrium.

Game theory raises many puzzles about rationality, which is why it is so fascinating to philosophers. Responding to the puzzles is a good way of learning about rationality. Colman insightfully reviews many of game theory's paradoxes and uses them to argue that, in games, people follow nonstandard principles of reasoning. He does not, however, claim that those nonstandard principles have normative force. Do principles of rationality need revision in light of the paradoxes of game theory? Rationality is a rich topic, and familiar principles are not likely to capture all its nuances. Nonetheless, standard principles of rationality are very versatile. In this commentary, I make a few general points about rationality and then show that extended applications of the standard principles resolve some paradoxes.

A standard principle of rationality is to maximize utility. The literature advances several interpretations of this principle. Nearly all acknowledge that impediments may provide good excuses for failing to meet it; the principle governs ideal cases. Game theory presents decision problems that are non-ideal in various ways. Perhaps in games some failures to maximize utility are excused. Rationality may lower standards in difficult cases.

In *Equilibrium and Rationality* (Weirich 1998), I generalize the principle of utility maximization to accommodate non-ideal cases in which no option maximizes utility. I assume that standards of rationality adjust to an agent's circumstances, so that in every decision problem some option is rational. The generalization of the decision principle leads to a generalization of Nash equilibrium that makes equilibrium exist more widely. A principle Colman calls "rational determinacy" supports the generalizations. The version I endorse asserts that rationality is attainable, but not that rationality is attainable in one way only. It allows for multiple solutions to a game. Associated with the principle of rational determinacy is the view that achieving an equilibrium is just one requirement of rationality, and meeting it is not sufficient for full rationality. Principles of rationality govern equilibrium selection also. I ascribe to the principle of equilibrium selection called pay-