Acta Psychologica 128 (2008) 409–412

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Team reasoning and collective rationality: Piercing the veil of obviousness

Andrew M. Colman [a,*], Briony D. Pulford [a], Jo Rose [b]

[a] School of Psychology, University of Leicester, Leicester LE1 7RH, UK
[b] School of Education, University of Exeter, Exeter EX1 2LU, UK

ARTICLE INFO

Article history:
Received 3 March 2008
Accepted 2 April 2008
Available online 19 May 2008

PsycINFO classification:
2340
3020

Keywords:
Cognitive hierarchy theory
Collective rationality
Evidential decision theory
Methodological individualism
Payoff transformation
Team reasoning

ABSTRACT

The experiments reported in our target article provide strong evidence of collective utility maximization, and the findings suggest that team reasoning should now be included among the social value orientations used in cognitive and social psychology. Evidential decision theory offers a possible alternative explanation for our results but fails to predict intuitively compelling strategy choices in simple games with asymmetric team-reasoning outcomes. Although many of our experimental participants evidently used team reasoning, some appear to have ignored the other players' expected strategy choices and used lower-level, nonstrategic forms of reasoning. Standard payoff transformations cannot explain the experimental findings, nor team reasoning in general, without an unrealistic assumption that players invariably reason nonstrategically.

> The veil of obviousness that makes so many insights of intuitive psychology invisible to our scientific eye has to be pierced. (Heider, 1958, p. 7)

1. Introduction

Collective rationality and team reasoning are common in everyday experience, but they have largely escaped the purview of game theory. To a layperson, it may seem obvious and natural that people frequently put the interests of their families, friends, and other groups to which they belong ahead of their individual self-interests. For decision theory and game theory, however, this has radical and subversive implications, because it challenges an entrenched though usually unarticulated and unexamined assumption of methodological individualism.

In cognitive and social psychology, collective rationality and team reasoning have been almost entirely ignored, although they represent a distinctive mode of reasoning and decision making and a quintessentially social psychological phenomenon. Our target article (Colman, Pulford, & Rose, 2008) reported experimental evidence that decision makers do indeed use team reasoning in certain types of games, and we are grateful to three distinguished commentators (Krueger, 2008; Sugden, 2008; Van Lange, 2008), who have commented on it with varying mixtures of agreement and disagreement. We reply to these three commentaries in the paragraphs that follow.

2. Evidential decision theory

Krueger (2008) agrees that orthodox game theory is "descriptively wrong" and that a psychological game theory is needed to account for interactive decision making, but he considers team reasoning's rejection of methodological individualism to be a step too far. He claims that evidential decision theory can adequately explain the relevant observations and experimental findings without abandoning the individualist foundations. Evidential decision theory rests on the assumption, in Krueger's words, that "most people have a strong expectation that members of their own groups will act as they themselves do". This provides a possible explanation for the intuitive appeal of the L strategy in the Hi-Lo matching game described in the target article (Colman et al., 2008, Fig. 1, right). Players who expect their co-players to act as they themselves do have a reason to choose L in this game, and to cooperate in the Prisoner's Dilemma game (Colman et al., 2008, Fig. 2, left), because these strategy choices maximize their payoffs, given the assumption that their co-players will choose the same strategies that they themselves choose. If a co-player is expected to mirror

* Corresponding author. Tel.: +44 (0)116 229 7197; fax: +44 (0)116 229 7196.
E-mail address: amc@le.ac.uk (A.M. Colman).

0001-6918/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.actpsy.2008.04.001

		II	
		B	*R*
I	*B*	0, 0	2, 2
	R	1, 1	0, 0

Fig. 1. The Ball Gown game, strategically identical to the Hi-Lo matching game.

		II	
		A	*B*
I	*A*	3, 3	2, 2
	B	4, 4	0, 0

Fig. 2. A game with a strictly dominant strategy for Player II.

any strategy choice, then *L* in the Hi-Lo game and cooperation in Prisoner's Dilemma are payoff-maximizing strategies.

Krueger criticizes the games used in our Experiment 1 on the ground that "if people care as much about the other's payoff as about their own, they will make the choice predicted by team reasoning (Van Lange, 1999)". In other words, the games used in Experiment 1 cannot distinguish between strategies chosen according to Van Lange's prosocial value orientation on the one hand and team reasoning on the other. Sugden (2008) makes a similar point, and we shall therefore put it to one side and return to it when we reply to Sugden below.

Turning to Experiment 2, Krueger claims that evidential decision theory can provide a satisfactory explanation for the findings. In that experiment, evidential decision theory and team reasoning predict the same choices, because the collectively rational team-reasoning strategies are also the strategies that maximize a player's payoffs against a co-player who automatically mirrors any strategy choice. If Players I and II always choose the same strategy, then the outcomes on the main diagonals of the payoff matrices in Fig. 5 of the target article are the only ones that can occur. On that assumption, ignoring all other cells, the payoff-maximizing strategy is *C* in every game. Krueger is therefore right: the dark bars in Fig. 6 of the target article could, in principle, represent choices based on evidential decision theory rather than team reasoning.

The obvious question that arises is whether this offers a more persuasive explanation than team reasoning. We believe that team reasoning is far more persuasive for several reasons. First, Krueger clearly considers evidential decision theory to be not only descriptively useful but normatively rational, for he takes issue with Hurley (2005) and Elster (1985) for suggesting that it is discredited. However, after it was introduced by Jeffrey (1965) and applied to games by Davis (1977), it was strenuously criticized by Lewis (1979, 1981) and many others. Most game theorists reject it for the obvious reason that players have no rational justification for assuming that others will act as they do. Even so, one might suggest that it is characteristic of the reasoning of boundedly rational decision makers, but everyday experience teaches that other people frequently behave differently from oneself.

A key assumption of Krueger's version of evidential decision theory is that a fixed probability can be assigned to Player II choosing the same strategy as Player I, *irrespective of which strategy Player I chooses*. This leads to absurdities, because both players have free will and make their decisions independently. For example, a player with consistent beliefs who assigns a subjective probability of .75 to a co-player choosing *L* in the Hi-Lo game necessarily assigns a probability of .25 to that co-player choosing *R*. The probabilities must sum to 1.00, because those are the only available options and one or the other must be chosen, and a player's own decision cannot affect the probability that the co-player will choose *L* or *R*, because the players are assumed to choose independently. But evidential decision theory assumes that players assign the *same* (high) probability to a co-player mirroring *any* strategy choice, and this seems incoherent.

Players with consistent beliefs cannot expect others to act as they do, independently of how they choose to act, because their decisions are not causally connected. In a strategic game, by definition, players choose independently. There is evidence (Anand, 1990; Quattrone & Tversky, 1984) that people sometimes lapse into forms of evidential reasoning in tricky dilemmas in which it is easy to be beguiled into believing that one's choices could influence the independent (in fact) choices of others. In our experiments, however, the games were straightforward, and the independence of the players' decisions was obvious and unambiguous, and for these reasons we do not believe that evidential reasoning can plausibly explain our results.

Another failing of evidential decision theory is that it predicts the same strategy choices as team reasoning only in games with symmetric team-reasoning outcomes. Consider the following game involving two women, I and II, preparing to go to a ball. Both own identical blue gowns (*B*) and identical red gowns (*R*), but I prefers *B* and II prefers *R*. Each has to choose which gown to wear to the ball. For each, the worst outcome would be to appear in the same gown as her rival, in which unhappy circumstances it would make no difference which gown she wore, and the best outcome would be to appear in her favorite gown while her rival wore the other. The payoff matrix is shown in Fig. 1. Each player is motivated to choose differently from the other, and the best for both is the top-right outcome (*B*, *R*), the collectively rational outcome in which each wears her favorite gown. Intuition and team reasoning predict this optimizing outcome as unambiguously as (*L*, *L*) in the Hi-Lo game, and it is obvious that the two women would choose it without any difficulty. But evidential decision theory suggests that either strategy is as bad as the other and is foredoomed to failure in any case because, as Krueger claims, "most people have a strong expectation that members of their own groups will act as they themselves do". This exposes a serious limitation of evidential decision theory as an alternative to team reasoning. What is more, the game shown in Fig. 1 is merely the Hi-Lo game with its columns transposed, and because the labeling of strategies is immaterial in game theory, we may simply say that it is the Hi-Lo game.

Before leaving Krueger's commentary, we need to correct a few errors and misapprehensions. First, it is not quite right to say that we do not entertain the possibility of evidential decision making. We refer to it in the second footnote in our target article, where we cite Colman and Bacharach's (1997) more circumspect version of it, which does not purport to be normatively rational. However, because Colman and Bacharach's theory never predicts a nonequilibrium outcome, and the team-reasoning outcomes in the experiments were all nonequilibrium outcomes, it cannot explain the findings in the target article.

Second, the following comment of Krueger's is based on a misunderstanding for which we may be to blame: "To guard against the reduction of collective preferences to individual ones, they suggest that the group may be treated as a 'singleton'. This is a remarkable set of ideas". What we said was that team reasoning reduces to orthodox decision theory when the group is a singleton. We meant merely that a group should be treated as a singleton when it is, in fact, a singleton (a one-person group).

Third, Krueger asserts that, in team reasoning, individual group members are not the decision makers: "They do not make decisions; the group does". As we presented it, team reasoning is a theory of individual decision making. An individual decision maker asks the question: *What do we want, and what should I do to help achieve it?* If group decision making were implied, the second part would be: *What should we do?* The decision makers are clearly individuals. Krueger's inference that team reasoning involves group decision making and is a throwback to discredited notions of the group mind is without foundation.

3. Cognitive hierarchy and level-*n* theory

Sugden (2008) agrees that the findings reported in the target article support the theory of team reasoning and that there is a strong case for including team reasoning among the social value orientations used in social psychology. He draws attention to a feature of the games used in Experiment 1 that limits our interpretation of its results, and he provides some supplementary analysis of the results of Experiment 2, suggesting that in addition to the team reasoning, some players may have used level-0 or level-1 individual reasoning.

The experimental games used in Experiment 1 were decomposable or separable, and we concede the point that they were therefore incapable of distinguishing prosocial strategy choices – maximizing the sum of the payoffs to both players – from team-reasoning choices. As already mentioned, Krueger (2008) has similar qualms about these games. The point is well taken, and Sugden's formalization sharpens and clarifies it, but it is worth commenting that our limited aim was to show that team reasoning predicts human choices better than game-theoretic Nash equilibria in certain games. We did not try to inoculate our experiments against all possible alternative explanations, and the prosocial value orientation provides one obvious alternative explanation in Experiment 1. Fortunately, the games used in Experiment 2 were not decomposable and were therefore immune to this alternative interpretation, although not entirely immune to evidential decision theory, as explained above.

Sugden's comments about cognitive hierarchy theory or level-*n* theory supplement our interpretation of the results of Experiment 2 and provide further insights into how the players appear to have reasoned about their decisions. Although Sugden's supplementary analysis is post hoc, his model is persuasive and consonant with the other research using cognitive hierarchy and level-*n* theory (summarized by Camerer, Ho, & Chong, 2004). According to the model, approximately half the players in each game used team reasoning, 40% used level-1 individual reasoning, maximizing their payoffs without regard to their co-players' payoffs – assuming, in effect, that their co-players were equally likely to choose any strategy – and 10% made random level-0 strategy choices, perhaps without any clear understanding of the games. The main objective of the experiment was to show that team reasoning would predict strategy choices better than game-theoretic Nash equilibria, and there was indeed very little evidence of Nash choices in these games.

4. Payoff transformations

Van Lange (2008) accepts the empirical evidence for team reasoning but suggests that it can be explained in terms of payoff transformations. We argued in our target article that payoff transformations cannot explain it, and Krueger and Sugden, in their commentaries, explicitly agree with this. We provided a formal proof that standard payoff transformations – linear transformations and the equality-seeking transformation, as clarified by Sugden – cannot explain team reasoning, hence Van Lange's claim appears incomprehensible at first. A closer examination reveals that his claim is, in fact, slightly different. What he shows is that an arbitrary nonlinear payoff transformation, coupled with a nonstrategic method of choice, generates predictions of team-reasoning strategy choices in a few of the games in Experiment 2.

The transformation labeled "MaxJoint and MinDiff (Integrative Model)" on the right-hand side of Van Lange's Table 1 includes an ad hoc addition of four units to each player's payoff in every outcome in which payoffs are exactly equal, and a subtraction of four units wherever the payoffs differ, irrespective of the degree of difference. Using Sugden's notation, Player I's utility is defined as $u_1 = x_1 + x_2 + h$ ($h = 4$ if $x_1 = x_2$; $h = -4$ if $x_1 \neq x_2$). After this transformation, the payoff matrix of Game 1 is as shown on the right of Van Lange's Table 1, ignoring for the moment the column labeled "Total". Only Player I's payoffs are displayed, but after transformation both players' payoffs are identical in every outcome, hence if both players choose *C*, the payoffs to Players I and II, respectively, are (20, 20), and so on.

The transformed game turns out to have no dominant strategies, and its Nash equilibria are (*C*, *C*), (*D*, *D*), and (*E*, *E*). Both players obviously prefer (*C*, *C*), where each receives the maximum possible (transformed) payoff of 20, and this is the team-reasoning outcome, but orthodox game theory provides no basis for choosing among the three strategies. Van Lange's method of strategy choice involves calculating the row sums, labeled "Total", and assuming that Player I will choose the strategy associated with the largest row sum. This implies entirely nonstrategic level-1 reasoning, taking no account whatsoever of Player II's expected strategy choices and assuming, in effect, that Player II is equally likely to choose any strategy. With these assumptions, the *C* strategy turns out to maximize u_1 for Player 1 and, by symmetry, u_2 for Player 2, and players who maximize u_1 and u_2 therefore choose the team-reasoning strategies.

The payoff transformation with the addition or subtraction of four payoff units is not only arbitrary but ad hoc. The standard prosocial payoff transformation, $u_1 = x_1 + x_2$, makes intuitive sense, representing as it does players who weight their co-players' payoffs equally to their own, but the further ±4 factor does not have any obvious intuitive or game-theoretic basis or justification, apart from producing the desired result in the game under analysis. Even so, it fails in some of the other games of Experiment 2. In Games 4 and 5 it generates *D* choices, and this highlights its strangeness, because *D* strategies are neither team-reasoning nor Nash equilibrium strategies, and there is no intelligible reason to choose them. Furthermore, if players choose their strategies according to Van Lange's nonstrategic method, by simply summing their own payoffs across rows or columns and then choosing the strategy with the largest sum, then the payoff transformation seems hardly necessary. If we apply this level-1 method of strategy choice to the untransformed payoffs in Fig. 5 of the target article, then players choose the *C* strategies in Games 3 (tied with *D*) and 4 – a result almost as successful as Van Lange's after his elaborate transformation.

It is instructive to examine Van Lange's method of strategy choice in the Hi-Lo matching game (Fig. 1, right, in the target article), because that game exposes the essence of team reasoning in its simplest form. After payoff transformation – summing the players' payoffs in each cell and adding 4 for payoff equality – Player I's row sum from choosing *L* is $(2 + 2 + 4) + (0 + 0 + 4) = 12$, and this is greater that the row sum from choosing *R*, namely $(0 + 0 + 4) + (1 + 1 + 4) = 10$, and similarly for Player II's column sums, hence both players choose *L*. But with this method of strategy choice, the transformation is once again unnecessary. Even without any transformation, it is immediately obvious that *L* yields a greater row or column sum of the player's own payoffs than *R*: $2 + 0 > 0 + 1$. This exposes the inadequacy of the whole approach because, as we explained in the target article, an individualistically rational player has no reason to choose *L* in this game, and this is universally acknowledged in game theory.

Van Lange's approach assumes that players invariably ignore their co-players' expected strategy choices and use only level-1 reasoning with transformed payoffs. But strategic reasoning performs most of the heavy lifting in game theory, and there is strong evidence from numerous experiments using many different types of games that "most players do *some* strategic thinking" (Camerer et al., 2004, p. 890, italics in original). In our Experiment 2, according to Sugden's model fitting (this issue), about half of our players' choices relied on a mode of (team) reasoning that requires attention to the co-player's expected choices. There are many games in which nonstrategic level-1 reasoning is evidently absurd. Consider the game shown in Fig. 2, for example. Player I, using Van Lange's method of choice, would note that (3 + 3 + 4) + (2 + 2 + 4) = 18 (from choosing row *A*) and that this is greater than (4 + 4 + 4) + (0 + 0 + 4) = 16 (from choosing row *B*) and would therefore choose *A*. But this would be very foolish indeed. It is easy for Player I to see that Player II is bound to choose column *A*, because column *A* ensures a better payoff for Player II than column *B* irrespective of Player I's choice – *A* is a strictly dominant strategy for Player II. Player I should therefore choose row *B*, to ensure a payoff of 4 rather than 3 when Player II chooses column *A*. This results in the best available payoff to both players, and it is also the team-reasoning outcome and the unique Nash equilibrium. Choosing strategies by simply summing payoffs (even transformed payoffs) across rows or columns and choosing the strategy with the largest sum is a hopelessly inadequate method. Interactive decision making requires strategic reasoning.

5. Conclusions

The results reported in our target article (this issue) provide strong evidence of collective preferences and team reasoning. Decision makers sometimes act to maximize the collective payoff of a group, and this cannot be explained in terms of standard social value orientations. Team reasoning is a distinctive and important mode of reasoning that should be acknowledged in cognitive psychology, and it should be added to the set of social value orientations used in social psychology.

The results of Experiment 1 can also be explained by the standard prosocial value orientation, according to which each player maximizes a utility function defined by the sum of both players' payoffs: $u_i = x_i + x_j$, but our primary objective was to discover whether players could be tempted to choose outcomes that deviated from unique Nash equilibria, and the results showed that most of them did. The results of Experiment 2 cannot be explained in terms of the prosocial value orientation, although evidential decision theory is not decisively ruled out. However, evidential decision theory fails completely in games with asymmetric team-reasoning outcomes that are as compelling as the symmetric team-reasoning outcome in the Hi-Lo matching game.

Most people are well acquainted with collective rationality and team reasoning from experience in everyday life, and many engage in it from time to time, but cognitive and social psychologists have so far remained almost completely oblivious of this elephant in the room. With strong experimental evidence for collective rationality and team reasoning now in the public domain, we hope that psychologists will begin to take account of these important phenomena.

Acknowledgements

This project was supported by Grant RES-000-23-0154 awarded by the Economic and Social Research Council of the UK to the first and second authors and a postgraduate studentship awarded by the Economic and Social Research Council of the UK to the third author. The first and second authors acknowledge the support given by their University in granting them study leave.

References

Anand, P. (1990). Two types of utility: An experimental investigation into the prevalence of causal and evidential utility maximisation. *Greek Economic Review, 12*, 58–74.
Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2004). A cognitive hierarchy model of games. *Quarterly Journal of Economics, 119*, 861–898.
Colman, A. M., & Bacharach, M. (1997). Payoff dominance and the Stackelberg heuristic. *Theory and Decision, 43*, 1–19.
Colman, A. M., Pulford, B. D., & Rose, J. (2008). Collective rationality in interactive decisions: Evidence for team reasoning. *Acta Psychologica*, doi:10.1016/j.actpsy.2007.08.003.
Davis, L. H. (1977). Prisoners, paradox, and rationality. *American Philosophical Quarterly, 14*, 319–327.
Elster, J. (1985). *Making sense of Marx*. Cambridge: Cambridge University Press.
Heider, F. (1958). *The psychology of interpersonal relations*. New York: Lawrence Erlbaum Associates.
Hurley, S. (2005). Social heuristics that make us smarter. *Philosophical Psychology, 18*, 585–612.
Jeffrey, R. C. (1965). *The logic of decision*. Chicago and London: University of Chicago Press.
Krueger, J. I. (2008). Methodological individualism in experimental games: Not so easily dismissed. *Acta Psychologica*, doi:10.1016/j.actpsy.2007.12.011.
Lewis, D. K. (1979). Prisoners' dilemma is a Newcomb problem. *Philosophy & Public Affairs, 8*, 235–240.
Lewis, D. K. (1981). Causal decision theory. *Australasian Journal of Philosophy, 59*, 5–30.
Quattrone, G. A., & Tversky, A. (1984). Causal versus diagnostic contingencies: On self-deception and the voter's illusion. *Journal of Personality and Social Psychology, 46*, 237–248.
Sugden, R. (2008). Nash equilibrium, team reasoning and cognitive hierarchy theory. *Acta Psychologica*, doi:10.1016/j.actpsy.2007.12.009.
Van Lange, P. A. M. (1999). The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of Personality and Social Psychology, 77*, 337–349.
Van Lange, P. A. M. (2008). Collective rationality: The integrative model explains it (as) well. *Acta Psychologica*, doi:10.1016/j.actpsy.2008.01.005.