

certain goal. (I argue for this at length in Gilbert 1989, especially pp. 167–203; see also Gilbert 1987.) If this is right, one can see how, from “We seek X,” plus appropriate premises about the means to achieve X, a conclusion about what an individual should do can follow directly, without the interposition of any assumptions about what *that individual* wants or seeks. Indeed, no single *individual’s* aims need be referred to.

Some further points, very briefly: It seems that part of the logic of “We seek such-and-such,” is that it implies a commitment not to act for the sake of personal benefit if this will prejudice “our goal.” In relation to a specified area of action, one has “given oneself over” to “our goal” (see Gilbert 1989, pp. 424–25).

A typical context in which “We” is used is a discussion: Participants are involved in a joint project (the discussion itself) and this may generate further joint projects. Given the logic of the corresponding “we” statements, it is clear that any joint project stands to generate a degree of self-abnegating and group-enhancing action.

The fact that folk psychology apparently allows for non-individualistic motivation, and the specifics of its concepts, should not be overlooked in the course of the scientific debate in which the authors of “Selfishness examined” are engaged.

The case of the “redundant” donor: Neither egoistic nor altruistic

Gene M. Heyman

Department of Psychology, Harvard University, Cambridge, MA 02138

The assumption that individual behavior is governed by rational self-interest has maintained a powerful hold on social science, psychology, and biology. In economics, it is usually assumed that individuals choose in order to maximize utility; in psychology, recent accounts of learning and motivation have proceeded from the assumption that organisms maximize reinforcement rate (or some similar variable); and in biology, genes have been depicted as “selfish” and rational according to the criterion of promoting self-replication. However, egoistically rational motivational theories have not gone unquestioned. It has been argued that such theories are incomplete because they ignore fundamental psychological factors, such as emotions (e.g., Frank 1988) and cognitive biases (e.g., Kahneman et al. 1982). Researchers who have compared matching law and optimization predictions have often concluded that matching is the more general principle and that when optimization does occur, it can be shown to be a special case of matching (e.g., Herrnstein 1981).

In the target article, Caporaël, Dawes, Orbell, and van de Kragt add their names to the list of critics who find that “economic man” provides too limited a model for human and subhuman motivation. They note that humans and other creatures often appear to behave altruistically rather than selfishly. Rational self-interest theories, which Caporaël et al. refer to as “egoistic,” account for this by trying to show that selfless acts, such as heroism and team play, are misleading, and that upon analysis such behavior proves to be as egoistic as the more transparent maneuverings of the entrepreneur. Caporaël et al. reject economic retranslations of altruism and offer instead a dualistic system. There are, they imply, two types of motives: selfish ones and altruistic ones. The problem, of course, is how to distinguish between them. Caporaël et al.’s solution is implied by the dichotomy. They designed experiments that they claim eliminated – subtracted out – egoistic motives so that the resulting behavior was necessarily altruistic.

Caporaël et al. ignore the possibility that human motivation is not exhaustively categorized by their two-part scheme: egoistic or altruistic. However, a large percentage of the subjects in the social-dilemma experiments described by Caporaël et al. made

decisions that appear to have been guided by a principle that was neither egoistic nor altruistic. These subjects, who comprised about one-half the sample, donated their initial dole, \$5.00, even though they stated that the donation was unlikely to help anyone else and they knew that it would be subtracted from any profits they might gain. Caporaël et al. label this sort of donation “redundant,” but they do not give the phenomenon special emphasis. Redundant contributions, however, as will be argued below, are theoretically important and call into question the apparently straightforward dichotomy between egoistic and altruistic motivations. A summary of some aspects of the experimental procedures and of Caporaël et al.’s approach will help show why this is so. Following this, it will be argued that redundant contributions were motivated by the desire for equitable social relations, and that this principle, equity, is neither simply altruistic nor egoistic.

In the standard procedure, each subject received a \$10.00 bonus if a critical number of fellow participants (e.g., 5 of 9) donated \$5.00. If less than the critical number donated, there was no bonus, and if more than the critical number donated, there was no additional benefit. Thus individual subjects who made what they believed to be a redundant donation absorbed a \$5.00 loss and to the best of their knowledge helped no one else. Put somewhat differently, a redundant contribution was by definition not egoistic (the \$5.00 loss), and the belief that the donation was redundant implies that the incentive could not have been altruistic.

Caporaël et al. explain redundant cooperators in terms of a sense of allegiance among players. According to a process referred to as “in-group biasing” and “group identification,” individuals react to themselves and others as exemplars of the group rather than as differentiated individuals. Caporaël et al. write (and this is the whole of their account of redundant donors): “Social identity, by reducing the distinctions between one’s own welfare and that of others, explains our subjects’ willingness to contribute in the face of expectations of their contribution’s being redundant.” Caporaël et al.’s explanation assumes that the motivation for redundant contributions was group welfare, but according to the pay-off structure, this could not have been the case – a redundant contribution, by definition, could not intentionally help anyone. One might argue that the subjects did not understand the social-dilemma contingencies; however, Caporaël et al. claim that the experimental procedure ensured that subjects clearly grasped the consequences of cooperating and defecting. Thus, the question remains, why did so many subjects contribute if they did not expect the contribution to enhance the winnings of the other players? A plausible answer is provided by equity theory. Caporaël et al.’s summary of egoistic theories suggests that they would label equity theory egoistic. However, below it is suggested that equity theory differs in significant ways from rational self-interest.

According to equity theory, social interactions are evaluated in terms of costs and benefits relative to the costs and benefits accruing to the other participants in the transaction (Brown 1986; Homans 1974). It is important to note that the scale is relative, not absolute. Thus, one prediction is that relative discrepancies rather than absolute levels of deprivation motivate action. Another, stronger, prediction is that under certain conditions (1) people can feel overcompensated and (2) they will make efforts to reduce personal undeserved advantage. Brown and Herrnstein (1975) cite an example. Adams and Jacobsen (1964) found that college students would work considerably harder at a proofreading task if they believed they were overpaid. The redundant donor in the Caporaël et al. studies may be a second such example.

Some subjects may have felt that it was not fair to earn more (\$15.00) for defecting (not donating) when others earned less (\$10.00) for cooperating. Unlike the college students in the Adams and Jacobsen study, the social-dilemma subjects did not

have the opportunity to invest more work, so the only way to reduce the discrepancy in equity was to reduce their pay. Thus, donations occurred even when they believed they would be of no assistance to the group. This interpretation could be checked by comparing the proportion of equity-based verbal reports from those who defected and those who donated with the understanding that a donation was redundant.

The assumptions associated with equity theory and the approach advocated by Caporaël et al. differ in interesting ways. Caporaël et al. assume that social interactions are motivated by either strictly egoistic or strictly social incentives. For example, their basic methodological strategy is to remove egoistic incentives and test whether cooperation persists. If it does, then by the logic of their strict dichotomy, it must have been maintained by social incentives. In contrast, the desire for fairness or justice, as depicted by equity theory, is synthetic in that it entails both egoistic and social ends. Individuals tally their costs and benefits and are goaded into action if the ratio seems discrepant. However, whether a cost-benefit ratio is acceptable is determined by how others are doing. Social mediation has two consequences: First, if all parties in the social exchange come to a similar judgment, the disparity, except for its sign, is equally felt. Second, an action that reduces a disparity in equity for one party necessarily reduces disparities in equity for all parties. For example, by working harder or taking a self-imposed pay-cut, an overcompensated employee can reduce the resentment that fellow workers are likely to feel. Thus, in contrast to the dichotomy of social and egoistic ends assumed by Caporaël et al., the desire for equity simultaneously serves both egoistic and group goals.

It should be pointed out, though, that equity may in some sense be a deficient or unstable equilibrium state. Note that the desire for equity does not necessarily lead to an optimal cost/benefit ratio, but simply the same ratio as others obtain. For example, if the participants in the social-dilemma experiments had not felt compelled to act fairly, they could have gone home at least \$5.00 richer, and done so without cost to others (assuming that their estimates of redundancy were more often correct than not, which appears to be true). Thus, if equitable relations are generally not optimal, it is surprising that they are not upended by more nearly optimal social exchanges.

Are redundant contributions a unique consequence of non-iterated social dilemmas or representative of a class of social motivations? The answer is not clear. On the one hand, the desire for equity is not the only incentive that promotes behaviors that enhance the welfare of more than one member of a social exchange. For example, good citizenship is rewarded and is in turn helpful for others. On the other hand, equity-based solutions are not necessarily optimal, and it is possible that equitable but nonoptimal social relations are unstable.

Selfishness reexamined: No man is an island

Alasdair I. Houston^a and William D. Hamilton^b

^aKing's College Research Centre, Cambridge University, Cambridge CB2 1ST, United Kingdom and ^bDepartment of Zoology, Oxford University, Oxford OX1 3PS, United Kingdom

Electronic mail: houston@vax.oxford.ac.uk

The target article combines experiments on human choice in a social setting with some general discussions on selfishness and cooperation in an evolutionary context. We find the experiments interesting and agree with the authors that they provide a challenge to sociobiological views of human behaviour. They do not seem, however, to imply such a radical difficulty for current sociobiological and evolutionary theory as Caporaël et al. suppose. In the first place, the extension of theory that they themselves suggest is needed is not really outside the existing framework. This is the major point discussed below. In the second place, the experiments, although throwing light on a

socially welcome aspect of human nature, may be no more puzzling, ultimately, than the maladaptive behaviour of a butterfly beating on a windowpane. In other words, psychologists' experiments, like other single-encounter social situations of civilised life, may be just as evolutionarily novel for us, and therefore just as confusing, as glass panes are to butterflies. This idea does not detract from the value of the experiments, which, if the idea is right, warn us that our behaviour is not stable at its present level of cooperativeness: Cheats, though uncommon, may be doing uncommonly well, and the proclivity to cooperate may be slowly declining (see also Hamilton 1971; 1975). Thus, the target article and its discussion may encourage us to consider remedial measures.

Turning now to Caporaël et al.'s own interpretations of their findings, one of our major objections concerns the characterisation of egoistic incentive (EI) theories and an apparently alternative Darwinian framework. Selfish gene theory is of much broader scope than Caporaël et al. suppose. It is in fact the basis of the current Darwinian approach. Although Darwin himself couldn't refer to genes, because they weren't known, the present theory has the same definite and individualistic spirit of interpretation that he used.

As follows from this, the authors' view of EI theory as part of sociobiology is open to objection. We clearly cannot speak for everyone who has used evolutionary considerations as a basis for speculation about human nature. To the extent that some people see selfish gene theory as a basis for only selfish behaviour, Caporaël et al. may be justified in the line that they take. We would argue, however, that their view of selfish gene theory and EI theories is misleading. They say: "According to EI theory, people will always choose the selfish strategy in social dilemmas" (sect. 1, para. 5). This is not true if kinship is involved or if there are repeated interactions. The authors go on to consider a variety of ways in which an individual's incentives may change in such a way that cooperation emerges. These ways are coercion, conscience, reciprocity, and inclusive fitness (IF) maximization. We feel that the last two categories do not really belong with the first two. The authors give the impression that IF theory attributes explicit incentives to individuals such that each rationally computes the IF associated with various options ("behavior . . . can be explained in terms of people's attempts to maximize their inclusive fitness" sect. 3, para. 12). Rationality is actually seen as a corollary of EI theory (sect. 1.4, para. 2). But IF says nothing about rationality, or the exact nature of the mechanistic and psychological processes involved.

Coercion and conscience may indeed change an individual's incentives so as to remove dilemmas, but reciprocity and IF theory do not rest on this basis. IF theory (Hamilton 1964) is a way of looking at the spread of genes, that is, a way of implementing the Darwinian approach that the authors themselves favour. Similar remarks can be made about reciprocity ("Darwin's emphasis on individual advantage has been formalized in terms of game theory. This establishes conditions under which cooperation based on reciprocity can evolve" [Axelrod & Hamilton 1981, p. 1,396].) The fact that such evolutionary conditions may underlie human behaviour does not necessarily remove a dilemma. We may in fact be conscious of a conflict between "our" interests and the maximization of IF.

Although we have argued that conscience differs from reciprocity plus IF as a potential explanation for cooperative behaviour, we are not barred from believing that conscience has been shaped by our evolutionary past so that certain sorts of behaviour are encouraged. (This form of evolutionary explanation may be an example of what Caporaël et al. refer to in section 1.4, paragraph 3 as "positing intervening selective processes." The use of such a two-level approach is by no means unique to cultural evolutionists: In discussions of the evolution of behaviour it is common to distinguish the evolutionary advantage and the behavioural mechanisms – see Houston 1980; Houston & McNamara 1988, for further discussion and references.) In