

CHAPTER 3

MOTIVES FOR UNCONDITIONAL COOPERATION

3.1 A first objection

Having pointed out the need for moral motives in the conventionalist theory of norms, we should ask ourselves, “Which moral motives?” In this chapter and the next we will take a closer look at the reasons agents might have for compliance with a norm in those situations where doing so is not straightforwardly rational. As we have seen this may be the case when the game is the last of an iteration, when non-compliance can be hidden from the group, or when the game is an iterated n -person prisoners’ dilemma. From now on, I will restrict the discussion to those situations where individual rationality precludes the stability of successful cooperation. This means that the attention will be focused primarily, but not exclusively, on the one-shot prisoners’ dilemma. In what follows I will refer to the reasons agents might have for compliance in situations where compliance is not straightforwardly rational, as *cooperative virtues*. Whether or not these reasons really are virtuous is a question I will take up later in this work.

As stated in the introduction to this study, I restrict the term *rationality* to the narrow and formal sense which most economists use. Rationality is characterized by the maximization of utility. A recurring theme throughout this and the next chapter is the opposition between the (proposed) cooperative virtues and considerations of rationality. There is nothing in the theory of rational choice that suggests that utility should be identified with self-interest. The utility function is simply the representation of the preferences of the actor over the alternative outcomes of his actions, whatever they may be. These preferences may express an altruistic, hedonistic or egoistic ordering. As long as this ordering of the alternative outcomes by the actor satisfies certain axioms – such as connectedness and transitivity – a utility function can be constructed representing (and predicting) his or her choice for action.

Bearing this in mind, it could be objected that I am approaching the question of cooperative virtues in the wrong way. If agents act cooperatively in a situation which I think is a prisoners’ dilemma, it probably was not one in the first place (assuming of course that people act

rationally in this minimal sense). Apparently these agents had preferences other than those of the prisoners' dilemma. Why not accept this and try to find out which preferences people really have?

This objection both misses and reformulates the point of this chapter and the next. It misses the point because it suggests that there is something fundamentally wrong with our intuitions about the interests people have. In normal circumstances, most people prefer more over less money, paying less taxes over paying more taxes, and so forth. Moreover, most people expect others to have these preferences in these circumstances. Contrary to the suggestion implicit in the objection, I do not believe we are so dead-wrong about these interests. If there is no kernel of truth to these expectations about the preferences of others, how is it possible that we have *these* expectations rather than another (or even none at all)? If this is so, we should not think of cooperative virtues as simply cooperative interests or preferences. Rather, cooperative virtues are attitudes or dispositions towards actions in games constituted by the interests people actually have. These dispositions prompt people to choose cooperatively in collective goods problems such as the prisoners' dilemma.

The criticism is also a reformulation of the task I have set for myself in this chapter. I want to find out what these dispositions towards one's actions are. What reasons do people have for not acting on their interests such that successful cooperation becomes possible? In what follows I will elucidate the sort of deliberations and considerations that characterize such agents.

3.2 Altruism

It is often believed that collective goods problems could be solved if only people were a little more altruistic. Proponents of this idea are right in locating the problematic character of the production of public goods in the untempered pursuit of interests by the relevant agents. The obvious reaction then is to hope people will consider the interests of others as they do their own and act altruistically.

What does altruism require? I will interpret altruism as the disposition that prompts an agent to consider the preferences of others as his own. These preferences give her reasons to act just as her own preferences do. Altruism, as I understand it, is indiscriminate with regard to the question whether those preferences are good for the agent concerned. Altruism is not a form of paternalism. The fact that others have preferences is what prompts the altruist to action, not the content of those preferences.

Howard Margolis as well as Nicholas Rescher have suggested models of the altruistic disposition.¹ Though these models diverge in their technical complexities, the basic idea is surprisingly similar. Altruists are thought of

¹ Margolis (1981), Rescher (1975).

as agents with (at least) two sets of preferences. One of these sets is their own set of ordered preferences, the other is the set of interests of (relevant) others.² These two sets generate a single utility function by some weighing device. Margolis thinks of it as an allocation function between self-interest and group-interest with diminishing marginal utility. Rescher employs a fixed rate of transfer, r , to get a resulting utility function. Both their models presuppose a measure of cardinal utility and the possibility of interpersonal comparison of utility.³

I will give an example using Rescher's idea of a fixed transfer rate. Suppose two altruists are engaged in the following game, the standard prisoners' dilemma:

		Column	
		C	D
Row	C	2 2	3 0
	D	0 3	1 1

Figure 3.1 the standard prisoners' dilemma

The payoffs in this matrix depict the altruists' interests. Suppose each has a rate of transfer of $r=1$, this means each considers the interests of the other as equally important as his own. After weighing the interests a new matrix is

² Margolis (1981) calls it the set of group-preferences. It is the ordering of alternatives as seen from the perspective of the entire group. According to Margolis, this implies that altruism extends to the entire group. Rescher (1975), on the other hand, allows for the possibility that one can be altruistic toward some and nonaltruistic toward others.

³ Both these assumptions are highly problematic from a theoretical point of view, especially, the interpersonal comparison of utility. I am in no position to specify exactly how this can be brought into the concept of utility. However there are some intuitive considerations which suggest that the assumption of interpersonal comparisons is not as awkward a notion as it is sometimes maintained. First, the reasons for action other people have are not completely private phenomena. Usually we have a good idea about what moves a particular person. For example, it is because we can imagine what we would experience in a similar situation. Secondly, because of this we can compare the interests and preferences of others with our own in terms of strength and importance. Admittedly, this does not amount to a justification of interpersonal comparisons of utility, but it does offer some intuitive background for both Rescher's and Margolis' assumptions.

rendered (see figure 3.2). Now both players have a dominant strategy: c (cooperate).

		Column	
		C	D
Row	C	4 4	3 3
	D	3 3	2 2

Figure 3.2 the transformed prisoners' dilemma

When an altruist deliberates about what to do, she first examines the possible outcomes of her actions. She determines the utility of each of these outcomes both to herself and to all concerned others. She then aggregates for each outcome her own utility and the weighted utility of the concerned others. Subsequently she selects the maximum of all the resulting aggregates and chooses the corresponding action.⁴

In many games, altruists will have no problem arriving at successful coordination. Compliance in those cases is less problematic for such agents than for similar agents lacking the disposition. Assurance games in the hands of two interacting altruists are reduced to impure coordination games in which the cooperative outcome is salient and risk-dominant (see figure 3.3). So altruists would comply with norms prescribing the optimal outcome, provided that the payoffs are known and have the right value and provided that the rate of transfer is sufficient. Of course the rate of transfer does not have to be $r=1$, or the same for everybody. In figure 3.3, Row might weigh Column's preferences with $r=1$ and those of Third with just $r=0.25$.

⁴ In formula: let U_s^i denote the utility to i of strategy s , where $s \in S_i$, the set of strategies available to i . Let r_j denote the altruist's rate of transfer connected to individual j . An altruist i chooses her strategy s such that $\max_i [U_s^i + \sum r_j U_s^j]$.

		Column			
		C	D		
Row	C	2	1	4	1
	D	0	1	1	2

$r=1$

Figure 3.3 altruists playing the assurance game

This seems a reasonable model of altruism. It captures the idea that altruists give weight to the interests of others in determining their own choice. However the model only says something about the sorts of choices altruists make. It has not really told us anything about the way altruists deliberate, except for the fact that altruists take into consideration the preferences of others.

There are two types of deliberation that can be subsumed under the category of altruism.⁵ One type I will call *sympathetic altruism*. This type of altruism does not imply a sacrifice of preference satisfaction. Let me illustrate by reconstructing the way such an altruist, Row, deliberates. First, she starts by identifying and ordering her own interests. She somehow finds out about those of Column. Next, she weighs the interests of Column and incorporates them into her own preferences. The satisfaction of Column's preferences has become satisfying for Row as well. The choice for cooperation is then dominant. Row's own satisfaction, which depends on that of Column, is maximally satisfied by cooperation. However, Row has not made a sacrifice in terms of her own preferences. At best, you could say that she has foregone the satisfaction of some of her preferences in favor of the maximal satisfaction of the whole of her preferences. The presence of the other with his specific interests and preferences has determined and changed Row's own preferences. Row does not experience a conflict of interests in the situation she finds herself.

This process of deliberation does not have to be a conscious, reflective process. It seems much more plausible that it proceeds through some form of instinct or reflex. This is not as strange as it may sound. For example, it seems paradigmatic of the way parents take pleasure in the well-being of their children, or the way lovers enjoy each others' delight. The idea has a respectable philosophical history. Spinoza argued that many of our emotions are the result of *imitatio*, the involuntary "mirroring" of the

⁵ The distinction that I present below is similar to the distinction that Schmidtz (1995, 98-102) makes between *concern* and *respect*.

<http://www.springer.com/978-1-4020-0639-5>

Instrumental Rationality and Moral Philosophy

An Essay on the Virtues of Cooperation

Verbeek, B.

2002, XII, 292 p., Hardcover

ISBN: 978-1-4020-0639-5