Economics and Philosophy, **17** (2001) 109–119 Copyright © Cambridge University Press

COLLECTIVE PREFERENCES, OBLIGATIONS, AND RATIONAL CHOICE

Margaret Gilbert
University of Connecticut

1. WHAT IS A COLLECTIVE PREFERENCE?

Can teams and other collectivities have preferences of their own, preferences that are not in some way reducible to the personal preferences of their members? In short, are *collective preferences* possible?[1]

In everyday life people speak easily of what *we* prefer, where what is at issue seems to be a collective preference. This is suggested by the acceptability of such remarks as 'My ideal walk would be . . . along rougher and less well-marked paths than we prefer as a family'.[2] One can imagine, indeed, that *each* member of a given family prefers something other than what the family prefers. What, then, do the collective preferences of everyday understanding amount to?

In his paper 'Team Preferences' Robert Sudgen argues that an account I have proposed is unsatisfactory, because it implies that when a collectivity has a preference, its members have associated *obligations*.[3] In this invited comment on Sudgen's paper I defend my account against this criticism.

I should emphasize at the outset that there are distinct species of obligation. To say that obligations are associated with collective preferences is not yet to specify which type is at issue.

[1] I prefer to write of 'collective' preferences rather than 'team' preferences (Robert Sugden's phrase) insofar as teams are a species of collectivity rather than the whole genus. Families, for instance, would not standardly be referred to as teams. The word 'collectivity' is rather cumbersome so I shall sometimes use the term 'group' for short.

[2] Sugden, 2000, p. 175. Sudgen writes here *in propria persona*. All references to Sugden in what follows are to this paper.

[3] Sugden, 2000, pp. 189–90.

https://doi.org/10.1017/S0266267101000177 Published online by Cambridge University Press

I should also stress that the focus of most of my discussion is the conception of collective preferences implicit in everyday discourse as opposed to any technical notion. When I refer to 'collective preferences' without qualification, this should be understood.

2. COLLECTIVE PREFERENCES AND OBLIGATIONS: IS THERE A CONNECTION?

i. Footballers

Do collective preferences involve obligations? In order to answer this it would be well carefully to consider some cases in which there is, by hypothesis, a collective preference.

Consider Sugden's preferred example of a collectivity, a football team. Assume that the team's objective is to maximize the probability that a goal is scored. Assume, further, that at a certain point in the game — call it t1 — this probability will be maximized if and only if team members Art and Ben both move to the right. The other possible combinations of an action of Art's with an action of Ben's all fare worse in this regard. All this is common knowledge between Art and Ben.[4]

I have not yet explicitly referred to a collective *preference*. Given that the team has a unique objective, however, it presumably follows it would rather reach that objective than fail to do so. More generally, a collective objective implies a preference for reaching that objective over failing to do so, absent any countervailing considerations.[5] In addition, given its specific objective, at t1 the team presumably prefers Art's moving to the right while Ben does to any other possible combinations of actions of these players.

Now suppose that at t1 Art moves left. We would expect Ben and indeed the other team members to be surprised. We would also expect their reactions to go beyond mere surprise.

Given that Art's moving left is not obviously explicable by some mitigating circumstance, his teammates will feel *aggrieved*. That is, they will understand that Art has *offended against them*. Either at the time or at some later point when conversation is more feasible they may say things like 'What do you think you are doing?' in a *punitive*, *rebuking* tone. Such rebukes are not likely to come from members of the opposing team.

[4] Many accounts of 'common knowledge' have been offered. Philosophers may be best acquainted with the pioneering discussion in Lewis, 1969. Economists may be more familiar with the independent work in Aumann, 1976. Informally one might say that the matter that is common knowledge is 'out in the open' with respect to the people in question. For some further discussion see Gilbert, 1989.

[5] A collectivity could of course have multiple objectives and rank them differently, where it was considered better to reach one while failing to reach another than vice versa.

https://doi.org/10.1017/S0266267101000177 Published online by Cambridge University Press

Members of that team might, indeed, express surprise after the fact. They are not likely to attempt to tell Art off.[6]

Whether or not rebukes are uttered by his teammates, Art will understand that *they have the standing to rebuke him*. He will understand, further, that there is a sense in which they were *entitled* to his moving to the right rather than to the left. Correlated with this entitlement is an obligation to them, an obligation to move to the right. His default on this obligation gives them the standing to tell him off. As to members of the opposing team, he had no such obligation to *them*, hence they lack the standing to rebuke him that his teammates have.

Why do the members of Art's team have this standing? Their opponents may have the same expectations as to what Art will do, and may rely on these expectations in deciding what to do themselves. Thus, though it may be tempting to do so, it is hard to argue that the basis of Art's fellow team members' standing is disappointed expectations or reliance.

The obligations and entitlements in question are grounded, rather, in the existence of Art's team's collective preference for his going to the right, a preference deriving from the team's objective. Thus those rebukes that are uttered may be backed up with such remarks as 'We're trying to win this game, not lose it!' and 'How are we going to win if you do things like that?' Indeed, such remarks may themselves be made as rebukes. If these observations about likely reactions and the understandings underlying them are granted, they must be accommodated by any adequate account of collective preferences.

ii. Conscripts and teenagers

Sugden takes two examples to cast doubt on the idea that there are obligations associated with collective preferences. These are the examples of a conscript and his unit and a teenager and her family. He writes that 'it is hard not to feel uneasy about the assertion that these individuals have obligations to participate in the plural agency of their groups'.[7]

Sudgen does not explain why he feels uneasy about these individuals. I suspect it is because a conscript has indeed been conscripted, with little choice as to whether or not to join his unit, while a teenager

[6] A member of the opposing team might possibly tell Art off if what he did amounted to ruining the game, thus making impossible the objective of the complex collective constituted by the two opposing teams, which is, one may assume, to play a decent game of football. He would then be making this rebuke on the basis of his membership in the complex collective as opposed to one of the opposing teams. I am assuming that Art's action, though flawed from the point of view of his team, does not amount to ruining the game in this sense, so a rebuke on behalf of the complex collective will not be in place.

[7] Sugden, 2000, p. 190.

https://doi.org/10.1017/S0266267101000177 Published online by Cambridge University Press

has little choice as to which family she is a member of. One may find it hard to accept that those entering a group in such circumstances thereby accrue obligations to promote the group's objectives and to act in accordance with any ensuing preferences.

In order to defuse this concern one might start by comparing the conscript and the teenager with the footballers. I shall take the case of the conscript first. In reality the situation of many conscripts is complicated by background circumstances involving some kind of consent to the practice of conscription, so I shall consider a case that does not involve such a complication.

Suppose that Jake is kidnapped and flown to a remote desert, where he finds many other men who, like himself, have been kidnapped and flown to the site. A loudspeaker announces to the assembled men that they constitute a military unit called the Desert Rats whose immediate objective is to advance five miles west on foot before sunrise.

It is unclear at this point that Jake is the member of a group with an objective, as opposed to a collection of individuals all in the same unfortunate situation. That there is such a group according to the announcement does not mean that there is one in fact.

Little may be needed to create such a group. For example, suppose Jake begins to talk without irony of *our* objective when referring to the objective ascribed to the Desert Rats. The other men who were kidnapped accept these references without demur. This may be enough to create the relevant collective objective and with it a collectivity comprising Jake and the others.[8]

Suppose, then, that in some such manner Jake and the others come to possess a collective objective. I propose that obligations corresponding to those of the footballers would then arise. Let us assume that, as is common knowledge among the men, in order to fulfil their collective objective, each must walk west without a break till sunrise. Should Jake decide to stop for twenty minutes, he will surely understand that the other group members have the standing to rebuke him for stopping. He will understand that they have this standing because they were entitled to his not stopping, or, to put it another way, he was obligated to them not to stop – by virtue of the collective objective.

Though Jake is a conscript of sorts, he is obligated to help promote the objectives of his unit. The circumstances that ground his obligation – his references without irony to our objective, and so on – are surely those of many conscripts.

The case of the teenager is similar. Suppose that Jane, a teenager, is speaking of her family. Without irony she refers to what we collectively prefer, and has such speech accepted as appropriate by the other family

[8] See the text below, Section 4.

https://doi.org/10.1017/S0266267101000177 Published online by Cambridge University Press

members. She will understand that now, if not before, they have the standing to rebuke her for failing to support the satisfaction of those preferences.

iii. Collective preferences as sufficient reasons for acting

In spite of the foregoing observations one may still be uncomfortable with the idea of a tight connection between collective preferences and obligations. The following concerns are suggested at various points in Sudgen's paper.[9]

Given that collective preferences are not reducible to a set of correlative individual preferences, it may be proposed that a *collective* preference as such cannot provide any of the *individuals* who make up the collectivity with a sufficient reason for acting.

I take it that if one has a sufficient reason to do something, then rationally speaking one *ought* to do it, absent any countervailing considerations. By 'rationally speaking' here I mean something like 'given that one is appropriately responsive to the applicable reasons'.

Accepting the proposal that collective preferences cannot provide sufficient reasons, one may infer that they cannot obligate, insofar as obligations are sufficient reasons for acting. A common form of reasoning suggests, however, that the proposal is false.

According to everyday understanding the following argument is valid: 'Our preference is to advance five miles west by sunrise. I, Jake, can best help to achieve our aim by resisting the urge to take a rest. Therefore I ought to resist that urge'. This argument goes from a premiss about (in Jake's terms) *our* preference to a conclusion about what *he* ought to do and hence has sufficient reason to do. There is no additional premiss stating a correlative personal preference of his own.[10]

I do not see how one can consistently accept the validity of this form of argument while maintaining that a collective preference does not provide group members with a sufficient reason for acting. It is true that, without more, it may be unclear precisely *how* collective preferences provide members with a sufficient reason for acting. But that is only to say that an explanation could usefully be forthcoming. I provide such an explanation shortly.

To say that collective preferences provide each member with a sufficient reason for acting is not to say that *the balance of reasons* must

[9] Sugden, 2000, pp. 189–90, 195. I may not always have understood Sugden's drift. In any case these are possible concerns that merit consideration.

[10] To my knowledge the intuitiveness of this form of reasoning was first pointed out by Wilfrid Sellars. See, for instance, Sellars, 1963. Having observed it independently I discussed it in Gilbert, 1989, and elsewhere, most recently in 1999 ms. See also Gilbert, 1994, and the text below.

https://doi.org/10.1017/S0266267101000177 Published online by Cambridge University Press

favor a given member's doing his best to satisfy the collective preference in a particular case. Rationally speaking, one may well be able to ignore the collective preference in the sense of not, at the end of the day, doing one's best to satisfy it.

One might think that were collective preferences to *obligate* one could not, rationally speaking, ignore them in this way. But this is not true of the obligations associated with collective preferences. According to everyday understanding, one may have a valid excuse for not fulfilling the obligation associated with a collective preference. Thus one might without logical oddity say to one's fellows: 'I'm sorry, but my conscience did not permit me to do it' where 'it' is what I needed to do in order that our objective be fulfilled. Note that here one implicitly grants that one's audience is entitled to an explanation and indeed an apology, though – at the same time – reason allowed one to act as one did.

4. THE PLURAL SUBJECT ACCOUNT OF COLLECTIVE PREFERENCES

The following key points have emerged in the foregoing discussion. First, collective preferences are not reducible to a set of correlative individual preferences. Second, they provide members of the group with sufficient reasons for acting. Third, they provide group members with obligations. More specifically, everyone is obligated to everyone else, an obligation violated if appropriate action is not performed. Thus each is entitled to appropriate action from the others, and entitled to rebuke those others for not acting appropriately. Fourth, these obligations are such that in certain circumstances a member can fail to fulfil them without irrationality.

My own account of collective preferences respects these points. I have been developing the details of this account since the mid-1980s.[11] I sketch it briefly here.

The core of the account is a concept of joint commitment. I take this to be fundamental for the understanding of everyday notions such as those of a collective preference.[12] A joint commitment is a commitment

[11] I first discussed collective preferences in Gilbert, 1987. This article focuses on collective belief of which it presents and defends an account, and argues for an analogous account of collective preferences that gives the lie to the common assumption among social choice theorists that 'no sense can be given to the notion of a social preference (or a group's preference) except in terms of (some function of) the personal preferences of the individuals concerned' (p. 211 in Gilbert, 1996). See also Gilbert, 1989, pp. 361–2, discussed briefly in Section 7 below.

[12] In Gilbert, 1987, I wrote of 'commitment as a body'. In Gilbert, 1989, p. 198f, I began to write of 'joint commitment'. Gilbert, 1996, Introduction, pulls together some key features of joint commitment as I understand it. Historically the use of such phrases as 'common will' and 'unity of wills' by political philosophers and contract theorists may well relate to this phenomenon.

https://doi.org/10.1017/S0266267101000177 Published online by Cambridge University Press

of two or more persons. In order to clarify what this amounts to I first say something about commitment in general and another species of commitment, namely personal commitment.

With respect to commitment in general, I focus on a particular example. I take a personal decision to do A to involve a commitment. That is, if I have decided to do A, I am subject to a commitment to do A. In contrast to this, inclinations and enthusiasms, for instance, do not yet involve commitments. I can be inclined to do A but not yet subject to a commitment to do A.

A personal decision can also serve to exemplify what I call a personal commitment. This is, by definition, the commitment of a single person who created it and who has the authority unilaterally to rescind it. One only has to change one's mind, as we say, and one is no longer committed.

It is now possible to clarify the concept of a joint commitment by distinguishing it from a set of personal commitments. A joint commitment is not a conjunction of personal commitments of the different parties. Each of the parties plays a part in the creation of a joint commitment, not by creating an appropriate personal commitment, but by expressing to the others his or her willingness to be jointly committed with them.

Each one understands that the parties will be jointly committed once all have made such expressions in conditions of common knowledge. Once in place, a joint commitment can only be rescinded by the parties acting together. I alone cannot change *our* mind; we can only do that together.[13]

Given that we understand ourselves to be jointly committed in some way, we do of course understand that each of us is subject to a commitment, but it is not a personal commitment in the sense defined above. I cannot rid myself of my commitment here just by changing my own mind.

As to its content, a joint commitment is a commitment to do something *as a body*, where 'do' is interpreted broadly enough to include beliefs and preferences among other things. The qualifier 'as a body' is important. To be jointly committed to believe something as a body, for instance, is not to be jointly committed *each to believe that p*. Rather it is to be jointly committed to constitute as far as is possible a single body that believes that p. I say that those who are jointly committed to do something as a body constitute a *plural subject*.

[13] Special side understandings may sometimes suggest that a joint commitment can be created or rescinded by just one of the parties. On such side understandings see, for instance, Gilbert, 2000, p. 67 n13. On the meaning of default, see Gilbert, 1996, pp. 14–16, 381–3, and elsewhere.

https://doi.org/10.1017/S0266267101000177 Published online by Cambridge University Press

As noted, mutual expressions of willingness to be jointly committed, in conditions of common knowledge, suffice for joint commitment. A standard way of expressing such willingness is the use, without irony, of the person plural pronoun as in Jake's 'Our objective is to . . .'.[14]

5. COLLECTIVE PREFERENCES AND OBLIGATIONS

I have argued at length elsewhere that the parties to a joint commitment are in an important and special sense obligated to one another.[15] For now the following remarks must suffice.

If I fail to act according to our joint commitment, you – as one of us – have the standing to rebuke me – as one of us. You have this standing because the commitment I have failed to honor is, precisely, a joint one. It is not a personal commitment. It is not mine and mine alone. Nor is it a conjunction of personal commitments, mine and yours. It is ours.

Your standing to rebuke may appropriately be said to be grounded in an *entitlement* or *right* to my acting in conformity with our commitment. I have a corresponding *obligation* to you, an obligation to conform. If I fail to fulfil this obligation, your right – which is a right *against me* – has thereby been violated, and you have the standing to rebuke me.

Of particular importance here is the fact that all this is so simply by virtue of what a joint commitment is. No moral principle need be appealed to. If you like, the obligations in question here are not moral obligations. They may be referred to as obligations of joint commitment to avoid misunderstanding.[16]

Surely conscripts and teenagers can have such obligations. As long as they express their willingness to be jointly committed, and thus enter a joint commitment, they will have the obligations that attach to all such commitments. *Particularly* when one has no option, one may be willing in the relevant sense. Clearly this is not a sense of 'willing' that implies this is something one would want outside one's particular circumstances.[17]

My plural subject account, then, entails that collective preferences provide group members with obligations of the appropriate sort, that is,

[14] See *On Social Facts*, Chapter 4, on the first person plural pronoun; also Gilbert 1999a, cited by Sugden.

[15] See in particular Gilbert, 1993, and, most recently and fully, Gilbert, 1999b.

[16] The obligations in question can be described as 'obligations to' a particular person or persons, or, as in the literature of deontic logic, 'directed obligations'. I am not arguing that all directed obligations are obligations of joint commitment. I argue rather that obligations of joint commitment constitute at least an important species of directed obligation, a species I conjecture to be associated with collective preferences and the like. For discussion of the contrast of the obligations of joint commitment with obligations of another important type see Gilbert, 1993.

[17] See Gilbert, 1993.

https://doi.org/10.1017/S0266267101000177 Published online by Cambridge University Press

obligations to a person or persons who have corresponding rights. At the same time it allows that collective preferences are not constituted by a set of personal preferences. As I argue below it also respects the observations that a collective preference provides members with sufficient reasons for acting, and that in some circumstances a member may without irrationality act contrary to it.

6. COLLECTIVE PREFERENCES AS PROVIDING SUFFICIENT REASONS AND ALLOWING FOR RATIONAL DEVIATION

With regard to the point about sufficient reasons, a joint commitment will have at least the reason-providing power of any commitment as such. In everyday thought, the personal commitment derived from a personal decision would normally be judged to make it the case that, all else being equal, one ought to conform to it rationally speaking. Thus, suppose I decide to go to France tomorrow, and do not subsequently change my mind. When tomorrow comes, however, I start off for Alaska. Given my decision, this action will seem questionable: I am apparently not doing what I rationally ought to do.[18]

With regard to the point about the possibility of rational deviation, this is something my account can accommodate. One can sometimes validly excuse oneself from violating a commitment, joint or not. What counts as a valid excuse in such a case, as in any other, will ultimately be a matter of judgement, that is, a matter of assessing the way in which reasons of various kinds play off one against the other.

7. COLLECTIVE PREFERENCES AND RATIONAL CHOICE THEORY

I have argued for a particular account of the collective preferences of everyday understanding, an account according to which such preferences involve obligations of a specified type. I take such preferences to be possible, and indeed to be prevalent, and importantly to affect the course of human life.

Though at the outset of 'Team Preferences' Sugden implicitly invokes the collective preferences of everyday understanding, informally discussing various preferences of his family as a group as opposed to the individual members, he doubts that such collective preferences involve obligations, and proposes a definition of 'team preferences' that does not have this consequence. Of course one is at liberty stipulatively to define a technical notion of team preference for theoretical purposes. It would

[18] On personal decisions as providing sufficient reasons see also Gilbert, 1993. There are clearly radically different *kinds* of reason for acting. That it would please my aunt Fay if I were to visit her is one kind of reason to visit her; that I have decided to visit her is a reason of another kind. This is clear enough without any theoretical characterization of the difference.

https://doi.org/10.1017/S0266267101000177 Published online by Cambridge University Press

be unfortunate, however, if acceptance of Sugden's definition led one to overlook collective preferences in the everyday sense.

In this concluding section I wish briefly to note two issues on which there is accord between Sugden and myself.

I argued in Gilbert, 1989, that given the existence of collective objectives in the everyday sense it evidently makes sense to talk, in the terms of rational choice theory, of a collective's *utility function*, where this is not logically derivable from the personal utility functions of the group's members and may differ qualitatively from any of them.[19] Insofar as one of the central aims of Sugden's 'Team Preferences' is to recommend the idea of such a collective utility function to rational choice theorists he evidently agrees with me on this point.[20]

Sugden sees the idea of a collective utility function as helpful to understanding human reasoning in a group context. He focuses on his 'footballers' problem' in which both of two particular players moving to the right rather than to the left is what will most likely lead to a goal for their team.

He first describes this situation as it might be represented in classical game theory. Each player – wanting his team to win – personally ranks their both moving to the right highest, both moving to the left next highest, and the other combinations of actions tie for worst outcome.

Sugden argues that were each player to look at the situation so represented in terms of maximizing personal utility, and ask 'what should I do?' there would be no conclusive answer. He would know he should do his part in the unique optimal outcome if his opposite number does, but he has no basis for inferring that his opposite number will. For his opposite number will have the knowledge he has, *mutatis mutandis*, and no more. I am not sure that there is indeed no conclusive answer in this particular case, but let that pass here. It is certainly true of many of the situations that game theorists refer to as coordination games or problems, of which this is one.[21]

Sugden then considers how the footballers' problem would be represented as a 'team-directed decision problem': 'There are still two players, each of whom has to [choose] between moving left and moving right, but now there is only one scale of preferences on which outcomes

[19] Gilbert, 1989, Chapter 6, Section, 6.3: 'On the limitations of game-theoretical approaches to social convention and social phenomena in general', pp. 361–3.

[20] See Sugden, 2000, p. 176.

[21] Against the assumption of various others, I have argued that rationality does not always dictate that a given agent does his part in a combination of actions that is uniquely optimal according to the utility functions of all the parties. See Gilbert, 1981 (pp. 136–8 in Gilbert, 1996) and Gilbert, 1990 (pp. 44–6 in Gilbert, 1996). On why I am not sure about Sugden's footballers' problem see Gilbert, 1981 (Gilbert, 1996, pp. 150–52). The case I discuss there is not the same, but similar reasoning may be plausible.

https://doi.org/10.1017/S0266267101000177 Published online by Cambridge University Press

are ranked, the scale of team-directed preferences'.[22] Sugden goes on to contrast the situation in which agents consider what will maximize collective utility with a standard-issue coordination problem.[23]

These points are consonant with the discussion in Gilbert, 1989, where I argue among other things that to consider 'What shall we do in order to reach our collective goal?' is not to participate in a coordination problem in the game-theorists' sense.[24]

I am glad that Sugden, a distinguished rational choice theorist, finds these ideas important. Precisely how they are best applied to the understanding of human actions is a further topic there is no space to engage with here. Suffice it to say that in any fine-grained analysis of actions in the context of collective preferences the obligating character of these phenomena will doubtless play a role.[25]

REFERENCES

Aumann, Robert. 1976. 'Agreeing to disagree'. *Annals of Statistics*, 4:16–28

Gilbert, Margaret. 1981. 'Game theory and *Convention*'. *Synthese*, 46:41–93. Reprinted in Gilbert, 1996

Gilbert, Margaret. 1987. 'Modeling collective belief'. *Synthese*, 73:185–204. Reprinted in Gilbert, 1996

Gilbert, Margaret. 1989. *On Social Facts*. Routledge

Gilbert, Margaret. 1990. 'Rationality, coordination, and convention'. *Synthese*, 84:1–21. Reprinted in Gilbert, 1996

Gilbert, Margaret. 1993. 'Agreements, coercion, and obligation'. *Ethics*, 103:679–706. Reprinted in Gilbert, 1996

Gilbert, Margaret. 1994. 'Me, you, and us: distinguishing egoism, altruism, and groupism'. *Behavioral and Brain Sciences*, 17:621–2

Gilbert, Margaret. 1996. *Living Together: Rationality, Sociality, and Obligation*. Rowman and Littlefield

Gilbert, Margaret. 1999a. 'Reconsidering the 'actual contract' theory of political obligation'. *Ethics*, 109:236–60. Reprinted in Gilbert, 2000

Gilbert, Margaret. 1999b. 'Joint commitment and obligation'. *Utilitas*, 11:143–63. Reprinted in Gilbert, 2000

Gilbert, Margaret. 1999 ms. 'Reasoning about Us'. Revised version, presented at an expert seminar on rationality and intentions, University of Amsterdam, 1999

Gilbert, Margaret. 2000. *Sociality and Responsibility: New Essays in Plural Subject Theory*. Rowman and Littlefield

Lewis, David. 1969. *Convention: A Philosophical Study*. Harvard University Press

Sellars, Wilfrid. 1963. 'Imperatives, Intentions, and the Logic of ''Ought''''. In *Morality and the Language of Conduct*. G. Nakhnikian and Hector-Neri Castaneda (eds.). Wayne State University Press

Sugden, Robert. 2000. 'Team preferences'. *Economics and Philosophy*, 20:175–204

[22] Quotations from Sugden, 2000, p. 193.
[23] Sugden, 2000, p. 193.
[24] Gilbert, 1989, pp. 361–3.
[25] See Gilbert, 1999 ms for more on this topic.

https://doi.org/10.1017/S0266267101000177 Published online by Cambridge University Press