Michael Taylor

Cooperation, Norms, and Moral Motivation*

Abstract: It has been said that norms can solve collective action problems. To endorse a norm is to hold a normative belief. This article insists that we try to isolate moral motivation – motivation by moral belief – as such, and that its existence cannot be taken for granted. Accepting the Humean view that belief alone cannot motivate, the article rejects the thesis that there is a necessary or conceptual connection between moral belief and motivation; it warns that in looking for motivational powers or effects of normative belief we must take care to rule out the possibility that the motivation is merely derived from existing desires; and it argues that deliberation and evaluation do not produce desires purely out of beliefs. These considerations are among the necessary preliminaries to getting clear about the role of 'social capital' in solving collective action problems.

1. Introduction

Why do people act in the common interest in situations where there are strong individual incentives not to? (Or, in brief: Why do people cooperate in collective action situations?) Some answers to this question provide explanations based wholly on self-interest or on rational egoism. Important among these, there is first the approach which has been explored in much recent noncooperative game theory. The key idea here is that under certain conditions each of a group of actors finds it in her interest to act in the common interest when all (or most, or enough) of the others in the group do. This could be because they believe themselves to be playing a Prisoners' Dilemma or Chicken supergame and their discount rates are low enough to make mutual conditional cooperation an equilibrium. There are problems with this approach (as game theorists themselves are well aware [see especially Kreps 1990a], notably the co-existence of many equilibria, but they are not a warrant for abandoning the approach altogether (as some sociologists recommend). They suggest rather that the game-theoretic analysis must be supplemented, perhaps with one or more of the other arguments I shall mention. For it will be an uncommon collective action in which some form of rational con-

* This paper was written while I was a Fellow at the Center for Advanced Study in the Behavioral Sciences, Stanford. I am most grateful to Center staff for their support and to the National Science Foundation for financial support (#BNS-8700864). I was helped also by conversations with Dan Gilbert, Bill Lycan, Jamie Mayerfeld, Philip Pettit, Bob Pollack, Sara Singleton, Michael Smith, Bill Talbott, Bob Van Dyke, and Red Watson.

Analyse & Kritik 15 (1993), S. 70-86

ditional cooperation is absent. Even a person with an identity of some strong commitment-constituted sort is unlikely to want to cooperate regardless of what others do.[1]

There is, secondly, the argument that people act in the common interest because cooperation and noncooperation will be met with sanctions – positive and/or negative, centralized and/or decentralized, economic and/or social – other than the sanctions that could be said to be implicit in the compliance or non-compliance of others (to obviate repetition of this qualification, I'll refer to such sanctions as extrinsic sanctions).[2]

Against these self-interest-based explanations it is often said, especially by sociologists, that cooperation (and much else) is explained by norms, such as Kantian norms, norms of fairness, or norms against free riding. It is also said (sometimes by the same people) that action, including action in the common interest in collective action situations, is in some sense expressive, expressive of a self or an identity, and it is therefore identity that explains action (see for example Benn 1979; Hollis 1977; Cohen 1985). In both cases, action is reckoned to be non-instrumental and of a kind that cannot be accounted for by rational-choice theory. What are we to make of these claims? Here I shall consider only the first.

We must begin by observing that if we want to assess this claim about norms it does not help to start by defining a norm as in part a pattern of behavior generally followed and/or by stipulating that the presence of sanctions for noncompliance is a part of what it means for something to be a norm. I propose instead that we start by saying that a norm is just a standard. Then we can ask questions about whether and why people meet or conform to it, and the possible answers to such questions include self-interest-based explanations of the kinds mentioned above, in which compliance with the norm is explained as the result of rational conditional compliance or of extrinsic sanctions for compliance or noncompliance. But in these rational-choice explanations of compliance, no motivational work is being done by the norms themselves, and there is no distinctive motivation of a normative kind at work.

One new role for social norms in helping to bring about cooperation is in coordinating the expectations of players in a repeated game. The problem typically confronting the players of these games is that, if one cooperative equilibrium exists, then many do, and the players must somehow coordinate their strategies so as to ensure that one of them is selected. (This is not of course a problem among small numbers of players who can communicate.) Social norms might provide the players with the correlated expectations that would enable them to overcome this problem (see Kreps 1990a; 1990b).

There is another possible distinctive role for norms in helping to bring about cooperation. If a person subscribes to a norm enjoining some form of cooperation

---

[1] This is something that is not understood by certain writers who are critical of rational choice theory and espouse an expressive account of rationality or an explanation of cooperation in terms of identity.

[2] On sanctions of this kind see Michael Taylor 1982.

(doing her fair share, not free riding, or whatever), she may believe she should cooperate, and somehow such normative belief provides her with a motivating reason to act or in some other way influences her behavior independently of the effects of sanctions. Possessing such a normative belief may be said to be in part constitutive of subscribing to a norm or a consequence or concomitant of it. Or we might want to say that it is part of what we mean when we say that a number of people share a norm that each of them has a corresponding normative belief. In any case, what interests me here is the possibility that normative belief per se (as distinct from any sanctions that might back up the belief) can motivate action.

Now normative beliefs or judgments – beliefs or judgments that we 'should' or 'ought to' do some action – are various. We can first distinguish prudential beliefs (where the 'ought' is hypothetical or conditional) and non-prudential beliefs (where the 'ought' is categorical). Normative beliefs of the first kind are conditional on the presence of a desire, as where my belief that I ought to (or should, or must) work harder is intended to mean only: "If I want to finish writing my book this year, then I must work harder", or where my judgment that you should visit me more often means only that I think you'd better do so if you want not to be struck from my last will and testament. These we call prudential beliefs or judgments.

Whatever else they are, moral judgments are not of this prudential sort. It is a mark of a moral judgment that it is thought to apply to an agent even when she has no desire that would be satisfied by acting in accordance with the judgment. That I would not withdraw my belief that you should refrain from gratuitously torturing animals if I discovered that such restraint would satisfy no desire of yours is part of what makes this a moral belief.

This does not mean of course that a person cannot have a prudential reason for acting in a way that (as it happens) is consistent with some moral belief that he holds. Undoubtedly there very often is such a reason (and this usually makes it very hard to be sure that non-prudential reasons or causes influenced the action). Moral beliefs are usually shared and in that case there are usually present, along with the moral beliefs, non-normative considerations which motivate the person to act consistently with the moral belief. We have already seen two possibilities here: the consideration that, failing such action, others too will fail to act similarly, and that extrinsic sanctions of some kind will be brought to bear. In neither case is a person's motivation engendered just by his beliefs, moral or otherwise. In particular, what I am calling moral motivation, or motivation resulting solely from holding a moral belief or making a moral judgment, is not involved in either case. If it is a fact that most people most of the time are motivated to do what they believe they should do, this does not demonstrate the reality of moral motivation.

Similar things can be said about the role claimed for identity in explaining action, including action in the common interest in collective action situations, and also about certain claims about the role of ideology, since ideology is sometimes defined in such a way that ideologically driven behavior is just a species of self-expressive behavior, expressive of an identity constituted by ideological commitments. For others, ideological behavior is seen as a type of behavior that is driven

by belief – strongly-held normative, possibly moral, beliefs; and if this is true, then what I have to say about moral motivation applies also to claims about ideologically motivated behavior.

Many economists and other rational choice theorists appear to believe that purely instrumental motivations, or even just motivations of the two kinds mentioned earlier (those arising from conditional cooperation and social sanctioning), are all we need to explain human behavior and that there is therefore no need to appeal to what I am calling moral motivation or to identity considerations. Some doubt that moral motivation, when properly isolated as such, exists. (This is of course not at all the same as doubting that there is moral behavior, which has many sources, most of them not involving moral motivation.) This is not my view, although I do believe that conditional cooperation and social sanctioning suffice to explain a wide range of behavior.

But if it is true that economists and most rational choice theorists typically fail to recognize moral motivation at all, it is also the case that there are very many sociologists and anthropologists who claim that norms and/or identities motivate and explain action but who fail to isolate normative and expressive motivation properly from other sorts of motivation or make any serious effort to test alternative explanations of norm-abiding and self-expressive behavior (such as the ones referred to earlier, involving conditional cooperation and social sanctioning).[3] Getting clear about what is or might be involved in normative behavior would also help us to refine the large but still very embryonic claims that have recently been made on behalf of 'social capital', a rather loose assortment of things, including social norms, that are thought to facilitate mutually beneficial exchanges of various kinds, exchanges that would not be consummated by rational egoists lacking such capital (see especially Coleman 1990, ch.12, and Putnam 1992).[4]

In trying to get clear about the issues at stake (which is all I shall do here), let us begin by noting that if normative beliefs are just beliefs in the ordinary sense of the word, then to claim that normative beliefs can motivate is of course to claim that beliefs can motivate and this is a very controversial claim. I concentrate on moral beliefs, for which the claim has been explicitly made. Most norms no doubt affect behavior either because they are backed by sanctions (especially informal social sanctions) or via their role in coordinating expectations; if something else is at work, if in particular normative belief per se affects behavior, we are most likely to find it in connection with moral belief. On some accounts, a moral norm just is a norm enjoining cooperation in a collective action situation (i.e., in a situation where the actors, though they share a common interest in cooperation, are

---

[3] Amitai Etzioni's recent book, *The Moral Dimension,* 1988, is typical. He cites much evidence which is supposed to support the view that moral motivation is widespread. But most of it is evidence of various sorts of altruism and for much of it there are obvious plausible alternative explanations not involving moral motivation. No serious effort is made to ascertain how much of the evidence could be explained by any of these alternatives; indeed, the alternatives are not even spelled out.

[4] Coleman credits Glenn Loury with introducing the term "social capital".

each impelled by narrow self-interest not to cooperate). If this is so, then moral norms are the only sort of norms I need consider, given that my interest here is in what can explain cooperation in precisely such situations; but I shall not assume the correctness of this account of morality.

In the next section I state what I take to be the problem faced by anyone wishing to make the claim that moral belief affects behavior. Then, in the following section, I ask whether moral beliefs are in fact beliefs in the ordinary sense of the word. The ground I traverse in the next three sections has been much plowed by philosophers but rarely even seen by social scientists and it is for the latter that I write, with the help of the former.

2. The Problem (for Humeans): Belief, Desire, and Action

Let us begin with a central plank of the Humean orthodoxy about desire, belief and action. Hume argued (in the Treatise, Bk. II, Pt. iii, sec. 3) that "reason alone can never be a motive to any action of the will" and again that "reason alone can never produce any action, or give rise to volition"; and later (Bk. III, Pt. i, sec. 1), referring to the earlier section, he said he had proved "that reason is perfectly inert, and can never either prevent or produce any action or affection". Reason, he said,

> "in a strict and philosophical sense, can have an influence on our conduct only after two ways: Either when it excites a passion by informing us of the existence of something which is a proper object of it; or when it discovers the connexion of causes and effects, so as to afford us means of exerting any passion." (Hume 1888, 413f., 458f.)

He is usually taken to mean by the arguments from which these famous lines are taken that cognitive states, including beliefs in the ordinary sense of the word (beliefs, that is, in propositions which can be true or false), are in themselves motivationally inert; alone, they cannot motivate action. Or again: beliefs and desires are quite different kinds of states, and both are needed to produce motivation to act.

If this is so, and if it is also the case that moral beliefs are just like ordinary beliefs, then it follows that moral beliefs are motivationally inert, or: holding a moral belief (or making a moral judgment) does not by itself produce any motivation to act in accordance with it.

Yet it is surely the case that our actions are, at least some of the time, influenced by our moral beliefs or judgments; in fact, it is surely true that most people, most of the time, have some inclination to do what they believe they should do (though not of course a sufficient or overriding motivation, a motivation that outweighs all other motivations).

Hume did not embrace such a conclusion, but that is because he did not accept one of its premises: he did not believe that moral beliefs are just like ordinary beliefs. Rather, he said, in effect, moral 'beliefs' have both cognitive and conative elements, beliefs about matters of fact and some sort of motivation. He was what is

now called a noncognitivist about our moral experiences. In this, too, he has had and still has many followers.

While accepting (like most social scientists, whether they know it or not) Hume's insistence on the separation of belief and desire and (with a qualification to be noted shortly) the motivational inertness of belief, I shall not follow Hume in his moral noncognitivism. Moral beliefs, I shall argue, are just like ordinary beliefs and are therefore motivationally inert: there is no necessary or conceptual connection between moral belief and motivation. This allows of course for contingent connections.

So while Hume's way of talking about belief and desire poses a problem (for us) about moral motivation, his account of moral belief closes it (to his satisfaction), more or less definitionally. We open it up again by rejecting this account of moral belief.

But we have also to remember that Hume's argument that "reason alone [i.e., without a 'passion,' a desire] can never produce any action", that (ordinary) belief by itself is motivationally inert, cannot, as he formulates it and as I have just formulated it, be right. For there is some reason to believe ('social identification' theory may provide an example), and no good reason not to believe, that a judgment or the forming or the coming-into-cognitive-salience of a belief may cause action in ways that seem not to involve the formation of an intention: desire seems to be absent from the proximate causation of the action (though of course we can always trivially find, i.e. impute, a motivation to 'explain' any action; see Snare 1991, 85-87). Hume's argument has to be taken (as it is now usually taken) to be about the causation or explanation of action in terms of motivating reasons. But a great deal of action is not done for reasons. A belief (say, the belief, perhaps tacit, that 'you are not in my group') may cause me to act (voluntarily and intentionally but perhaps unthinkingly) in a certain way, though that belief may not have been part of a motivating reason (the other part of which must be a desire). Hume's argument does not apply to such cases. It applies only to cases where the agent acts for reasons, and in these cases a desire (using the term broadly) must be present as well as belief. (The desire which must be present in a motivating reason need not be one that the agent explicitly takes account of;[5] it need not be one that the agent recognizes himself as having in his deliberations, one of which he is consciously aware; it need only be present 'in the background', a mere disposition to act, and not necessarily 'in the foreground' of actual deliberation.[6])

Now it seems to me that among the cases it does not apply to are cases where moral beliefs have a causal influence on action in the absence of desires. Of the moral behavior that is not the product of some form of rational conditional coop-

---

5 This point is noted by Pettit 1987.

6 For the foreground/background distinction, see Pettit/Smith 1990.

eration or social sanctioning (see sec. 1 above), much is surely of this kind: it is caused by (possibly tacit) moral belief but not done for (motivating) reasons.[7]

3. Moral Beliefs are Beliefs

Our view of the motivational consequences or concomitants of moral belief clearly depends on what we take moral belief itself to be. Are moral beliefs beliefs about matters of fact? Do they have as their objects propositions that are true or false? In other words, are moral beliefs like beliefs in the ordinary sense of the word? When you claim, or state a belief, that abortion is wrong, or that rape is wrong, or that it is wrong to refuse to participate in a system of reciprocity from which you will benefit whatever you do, are you stating propositions whose truth is independent of what people happen to feel or believe (a 'realist' position)? Or are you saying only that these are things of which you disapprove (an 'emotivist' position)? Or that the claims would be true if and only if everyone in your society or group, or a majority of it, believed them (a 'constructivist' position)? There are really three broad positions on this (exemplified in the three answers just given); or rather, views on the nature of moral belief are parts of larger views on the nature of morality which fall into three groups. All these views are controversial.

There is first the objectivist view of ethics known as moral realism according to which it is the case that there are moral facts and true moral propositions (moral judgments being statements of these facts or truths) and that these facts or truths are independent of the beliefs that are our evidence for them.

The opponent of moral realism must therefore either deny that there are moral facts or true moral propositions or accept the existence of moral facts or truths but deny that these are independent of our beliefs. The first route is taken by noncognitivists, who claim that our moral experiences must include a non-cognitive part: they cannot consist only of beliefs about matters of fact, without any affective or evaluative response to these facts. Noncognitivism, which in moral philosophy proper has been the chief rival of moral realism, comes in various forms. One well-known version is emotivism, according to which moral terms are expressive in meaning and moral judgments are expressions of approval, of pro-attitudes, or feelings. Another version is prescriptivism, according to which moral terms are prescriptive rather than descriptive in meaning and moral judgments express recommendations or prescriptions.

The other contender against moral realism is constructivism, sometimes referred to as idealism. Constructivists allow that there are moral facts or truths but contend that these moral facts or truths are constituted by our moral beliefs. Most social scientists and political theorists will have encountered variants of one or both of two forms of moral constructivism. In the relativist form, variants of

---

[7] For further discussion of Hume's and Humean views, see Smith 1987; Pettit 1987; de Bretton Platts 1991, esp. ch. 4; and especially Snare 1991. I have benefitted from all of these in writing this and the following two sections.

which are favored by some anthropologists and sociologists, "a moral claim x (e.g., that abortion is wrong) is a moral fact for S (x is true for S) just in case S believes x, S would believe x upon reflection, S is a part of a social group the majority of whom believe x, or some such thing" (Brink 1989, 20). According to the nonrelativist version, held by a number of political philosophers (including Dworkin, Rawls, and Habermas), "there is a single set of moral facts that are constituted by some function of our moral beliefs, often by our moral beliefs in some favorable or idealized epistemic conditions" (Brink 1989, 20).

So constructivism shares with moral realism the view that there are moral facts or truths, and it is this, their shared "moral factism" (Copp 1991) which is all that I want to defend here. I leave undecided (because I do not here need to decide) the truth of the remaining part of constructivsm, which claims in effect that we construct moral facts. But we should note in passing that if a society, or any group of persons, views moral claims in certain highly relativistic ways – if in particular the view is widespread in that group that a moral claim is a moral fact for each person just as long as he or she believes it – and if there is wide disagreement among the members of the group about moral truth, then we might not see in that group effective practices of moral education, including the social shaping of moral emotions, which are essential for one sort of moral motivation.

If the noncognitivists are right, then it follows either that no good sense can be made of the question of whether moral beliefs motivate (because, for example, strictly cognitive and conative elements cannot be conceptually isolated within 'moral judgment') or that to have a moral belief is ipso facto to be motivated. Noncognitivists usually draw the second conclusion. Or, like Hume, they derive their noncognitivism from what they take to be the facts that moral considerations do guide our actions while no purely cognitive state, like a belief, could so motivate. (In Hume's words, " ... morals ... have an influence on the actions and affections," but "reason alone can never ... have such influence"; Hume 1888, 457.) Either way, noncognitivists are internalists about morality and motivation. They believe that we cannot make detached moral judgments; we cannot make a moral judgment about something without also having an affective or evaluative response to it. They hold that to be motivated – to hold a pro-attitude to what is judged right, fair, just, etc. – is a part of what it means to hold a moral belief or to make a moral judgment. One who denies this – who denies that the connection between moral belief and motivation is a necessary or conceptual or internal one – is an externalist about morality and motivation. The externalist believes that a moral belief alone will not constitute a reason to act but must be joined by a desire of some sort if the agent is to have a motivating reason to act.

If the noncognitivists are right, then we do not need to trouble ourselves with the problem of moral motivation. But noncognitivism is not uncontroversial, and I side with those who would reject it. It is at odds with the ways in which we ordinarily think and talk about morality. We think and talk about morality just as if moral factism were true; i.e., as if it were the case that there are moral facts and true moral propositions and that moral beliefs are like ordinary beliefs. We deliberate and argue about what is right and wrong, good and bad, just and unjust,

about what our moral obligations are, and so on, as if things (actions, policies, persons, etc.) have moral properties, as if there are ascertainable moral facts and assessable moral truths. We talk and write, adducing reasons and evidence, as if people can be wrong or mistaken about their moral claims. We think we can correct, and often we try to correct, our own and others' moral beliefs. We think and speak as if there are moral truths to be discovered. We recognize and seek to persuade others to recognize moral facts (compare Brink 1989, 23-39; Sturgeon 1986).

This is not the way we think about any sort of judgment: about aesthetic judgment, for example. We are willing to admit that our judgments about, say, music are in the end matters of taste. I know that there is no point in trying to persuade you that Schütz is better than Schoenberg, or at any rate that reasoned argument will not get me very far, if you have the opposite preference.[8] But most of us strongly resist the idea that our moral judgments are matters merely of taste or preference (compare Smith 1989).

Why would we bother to engage in all these argumentative and deliberative activities, as we do, if we thought there were no moral truths? Much of this behavior simply would not make sense if those who engaged in it denied that their claims could be true.

We also often think that we should be guided by these moral truths, that there exist, independently of our own inclinations and desires, moral facts by which we should be constrained. Our beliefs, said Frank Ramsey, are "maps by which we negotiate our way in the world" (Ramsey 1978, 134), and it would seem (contra Ramsey) that moral beliefs often function for us in the same way.

Our deliberative and argumentative behaviors, then, do not make sense if the noncognitivists are right, and so I shall proceed to the next section with the understanding that there are moral facts and true moral propositions and that moral beliefs are like beliefs in the ordinary sense of the word.

4. Against Internalism

Morality is usually taken to be action-guiding: we engage in deliberation and argument about moral matters and we offer moral advice because we want to provide ourselves and our interlocutors with reasons (preferably motivating) for action; we expect our conclusions and our advice to guide our own and others' actions (compare Brink 1989; Smith 1989). This action-guiding character of morality is of course captured by, is indeed built into, the account given by moral non-cognitivists of moral terms themselves, an account which is at once irrealist about moral facts and internalist about the relation between moral belief and motivation: moral judgments are not purely cognitive but express pro-attitudes or feelings. The action-guiding character of morality also drives some moral realists to internalist positions (using arguments which, though ingenious, have not found

---

[8] In which case you probably can't tell Scheidt from Schein.

wide acceptance – as, for example, in the work of Foot 1978, chs. XI and X, Nagel 1970, and McDowell 1978). But moral factism (and therefore also moral realism) does not entail or require internalism, as some moral philosophers (for example Nagel 1970; Mackie 1977) seem to suppose, and having rejected non-cognitivism I wish nevertheless to defend externalism in the following way. Though moral beliefs are just like ordinary beliefs and in themselves are motivationally inert, morality is 'normally' action-guiding: people are 'normally' but not necessarily motivated to act in accordance with their moral beliefs – the relation between belief and motivation is contingent, not conceptual or internal. Just as moral non-cognitivism seems not to fit the ways in which we deliberate and argue about moral matters, so also, it seems to me, does internalism sit uneasily with other facts about our moral experiences.

The internalism I wish to reject claims that it is a part of the very meaning of morality – it is a conceptual truth about morality (as opposed to a truth contingent upon what the moral judgment is or upon what [other] attitudes the agent has) – that moral belief or judgment (as distinct from moral obligation or the agent's recognition of a moral obligation) necessarily provides some (as opposed to sufficient or overriding) motivation (as opposed to justifying reasons for action, which may not motivate).[9]

An internalist of this sort (henceforth: an internalist) must show that there cannot be a person who judges or believes that he should do something yet has no desire, feels no inclination to do it, a person who recognizes that a moral fact applies to him but is indifferent to it, who possesses a moral belief that is motivationally inert.

But surely some people, if only a very small number of them, or perhaps many people some of the time, do seem to make moral judgments yet remain unmoved by them? The psychopath – or a certain sort of psychopath – is standardly referred to in support of this claim.

To this the internalist can reply that such a person, a person who is wholly unmoved by a moral belief he is supposed to hold, does not truly or genuinely hold the belief. He may recognize that it is a belief that most others hold, perhaps also that he should have this belief, and he may claim that he subscribes to this belief and may think of himself as holding it; but he does not in fact hold it.

Moral philosophers who have taken positions in this dispute have not, as far as I can see, produced a single knock-down argument or any persuasive set of arguments for or against internalism; and neither they nor anyone else that I'm aware of has produced convincing evidence.[10] I can do no better. It therefore seems right not to take the easy path by assuming a necessary connection between moral judgment and motivation. If I am wrong in coming down on the externalist side,

---

9 In Brink's terms (1989, 40-42), this is weak agent internalism about motives.

10 " ... we may indeed have empirical grounds for supposing that morality moves us to action *in some way or other*, but we do not have empirical grounds for supposing that morality - *alone* influences us. Are we sure that no further desire or passion was required such as a desire to be moral?" (Snare 1991, 55; citing Harrison, 1976, 1).

then of course I would not have had to worry about whether making moral judgments produces any motivation at all. In particular, internalism would have guaranteed that moral beliefs enjoining cooperation would be a source of motivation additional to the sources canvassed earlier (in sec. 1). But the need to study the separate effect on motivation of moral belief per se would remain. For the stronger internalist position – that to have a moral belief is necessarily to possess sufficient or overriding motivation to act – is obviously untenable; so that, even if some motivation is conceptually guaranteed by moral judgment, we would still want to know what could strengthen the connection between belief and motivation or produce additional connections.

Having rejected internalism, we are left with the conclusion that either there is no connection at all between an agent's moral belief and her motivational state or there is a contingent connection – contingent upon the content of the moral belief or upon facts about the agent other than her possession of the moral belief in question. The latter is the only plausible conclusion. That the relation between moral belief and motivation is not necessary or conceptual but contingent allows of course for the possibility that it is very common for moral belief to have motivation as a concomitant or consequence and for the connection to be not fortuitous but regular and explicable.

It surely is the case that most people most of the time have some motivation to do what they believe they should do. Of course this could be a wholly spurious correlation. More plausibly, it could be that in very many situations where a person has a moral belief there are usually also present non-normative considerations which motivate the person to act, as it happens, consistently with her moral belief – for example, the consideration that, failing such action, social sanctions will be brought to bear or that mutually advantageous cooperation will collapse. More generally it could be that the motivation is simply derived from the moral belief and a pre-existing desire (together, perhaps, with other desires and non-moral beliefs). Derived desires are the subject of the next section.

5. Derived Desires, Evaluation, and Second-Order Desires

Many of our desires are derived desires, being derived from our other existing desires together with beliefs. Example: Margaret Thatcher wanted to win re-election; she believed that winning the war over the Falklands/Malvinas would help her do that and she believed that sinking the Belgrano would scupper the peace negotiations in progress; so she had a (derived) desire to sink the Belgrano. Or so it has maliciously been said. The derivation of a desire may be immediate and unconscious; or it may require extended reasoning. In any case, the 'new' wants, if that is what they should be called, are implicit in existing desires and belief, and whether we say that the derivation involved 'deliberation' or something else, it does not derive desires from beliefs (or judgments, or evaluations) only.

And so it may be with motivation that appears to have been generated solely by a normative belief. A person believes that to ø is wrong and that therefore she should not ø. We observe that she also desires not to ø. A possible source of this

desire is a (pre)existing desire: some version, perhaps, of a general desire to do the right thing, itself having various possible sources.[11] From this desire together with her belief about ø is derived her desire not to ø. Of course, her desire not to ø may be derived from an existing desire and a non-moral belief. She may, for example, desire to maintain her self-esteem or to avoid disagreeable feelings, and believe that ø-ing will damage her self-esteem or trigger a disagreeable feeling. In none of these cases is a desire engendered solely by a normative or any other belief.

Of course, when extrinsic sanctions (of any of the kinds mentioned in sec. 1) are present, the desire to do what one believes one ought to do may just be of this derived sort: derived from a desire to avoid the sanction together with a belief that not doing what one ought to do will incur the sanction.

There is nothing mysterious here, but having made this simple point about derived desires, we can go on to consider a claim according to which a person can in effect produce desires by a purely cognitive activity.

We do of course sometimes evaluate the desires we already have, not only in the sense that we think about just how strong each desire is (just how much we really want something), about the competition and conflicts among our desires and of how we might satisfy some of them, and so on (Charles Taylor calls this "weak evaluation"; Taylor, C. 1985; 1989, esp. ch.1), but also in the 'strong' sense that we think about which of our desires are worthy desires, which are desires we ought to satisfy or to suppress, and so on.

But it is sometimes further claimed that we can form a new desire or suppress an existing one or make one of our desires the motivationally effective one purely by the cognitive processes of deliberation and evaluation, deliberation and evaluation informed perhaps by our normative beliefs. This I wish to deny. Consider an example, one chosen by a philosopher who has recently pressed the anti-Humean case on this point in some detail, mostly through discussion of examples like this. A woman and her partner have the opportunity of adopting a child in need of foster parents. Let us suppose that she has not thought about such a possibility before and she has no ready-made preferences for or against fostering this child, or any child. She deliberates, thinking of the good she can do and of the various ways in which fostering the child will affect her and others connected to her, and she evaluates the different motivations or desires that each of these considerations gives rise to, not only weighing the advantages and disadvantages, but trying to form judgments about whether each consideration should have any weight, should count as a reason at all. And eventually she 'decides' that she wants to foster the child.

---

[11] Compare the lines quoted in the previous footnote. It may be that the desire to do the right thing, to avoid wrongdoing *as such*, is a widespread desire, even if it is generally a weak one. But this of course does not obviate the need to distinguish the motivation derived from this (or any other) desire from moral motivation, i.e. motivation by moral belief.

The example is David Milligan's and Milligan wants to conclude from this and his other examples that there are desires that are "formed by the agent, chosen by him, and often the result of evaluation and deliberation" (Milligan 1980, 46). But there is nothing in Milligan's discussions of these examples to warrant a conclusion of this kind. We are indeed often faced with opportunities for choice among alternatives over which we have no preferences. In such cases, if we do not simply ignore them, declining to make a choice or to form a preference, we might initially choose one of the new courses of action without forming a preference (pace revealed preference theory). Or, like the woman in the example above, we might deliberate about the alternatives and so come to form (more or less considered) preferences between them. But might that not be because we have not only a motivation to propel us into doing this but also a desire or desires from which (together with our beliefs) the final preference is desired? So it is in Milligan's example: he assumes that the woman wants to help a child in need (and believes she can help this one), wants to continue with her drama work with a group of disabled adolescents as well as her own acting (and believes these activities will have to be curtailed if she adopts the child), and so on. There might also (Milligan doesn't say) have been some motivation arising from the emotion of shame (about which see below). We might also suppose – Milligan doesn't explain - that she was motivated to deliberate in the first place – rather than ignoring this choice and not trying to form a preference – by a desire to do the right thing and her belief that turning away from the choice would not be to do the right thing, or by her desire to avoid the disapproval of those who knew of her opportunity, etc., or again by the motivation arising from an emotion of shame. There is nothing in Milligan's discussion to show that her deliberation and her evaluation of her desires produced motivation beyond that produced by derivation from existing desires and beliefs (and perhaps directly by emotion), nothing to show that any part of her final preference was produced out of deliberational and evaluational thin air.

Evaluation, at least of the sort I've described here, does not, then, produce desires out of beliefs, normative or other, or produce desires by some purely cognitive process.

We might be tempted to think that evaluation produces a special sort of desire, namely a second-order desire, a desire to possess a particular first-order desire or pattern of first-order desires or a desire for a particular first-order desire to be extinguished or again a desire that a particular first-order desire should be the motivationally effective one, the one that moves the agent to action. (Second-order desires of this last kind are called second-order volitions by Harry Frankfurt 1971.) We might be tempted to think, in other words, that if an evaluation yields the judgement that I should have a certain desire, then if this does not produce the desire it should at least produce a desire to have the desire. Or, if evaluation yields the judgment that, of the desires I have, one of them, D say, should be the only motivationally effective one for me (or, as Frankfurt says, should be my will), then if this judgment does not itself bring it about that D is the only motivationally effective desire, it should at least produce a desire for this to be so.

We should resist this temptation (though some have not). Second-order desires are still desires and the argument against a necessary connection between judgments and desires applies to them as to any other desires. Arguments similar to those used against the case for a necessary connection between moral belief and motivation can be used here. Consider the case of the addicted smoker. She wants to smoke; but if she also wants her health and believes that smoking is detrimental to it, or wants the approval of her friends and believes that it would be greater if she were not to smoke, or whatever, then she'll also have the (derived) desire not to smoke. (Why, indeed, would she have the belief that smoking is detrimental to her health? After all, there is an infinity of beliefs she could hold but does not. Why this belief? Why does she even pay any attention to the subject of this belief? Presumably because she cares about her health.) From what standpoint, or on what ground, might she evaluate these two conflicting desires? There may be none, and so she may not judge the two desires, but just live uneasily with both of them, being moved to action simply by the stronger of the two. Of course, if her desire for health or for the approval of her friends is very great, so too will be her derived desire not to smoke, which may then very often be great enough to overcome her desire to smoke. It may also be the case that her desire to protect her health or her desire for approval, etc., are so strong that, recognizing her occasional weaknesses, she may wish that she didn't want to smoke or that her desire not to smoke was the only one ever to move her. But this second-order desire is not in this case the product or concomitant of an evaluation; it too is derived from her beliefs about the effects of smoking and a desire to protect her health, etc.

In this case, a second-order desire is formed without the prior or concomitant formation of an evaluation. When evaluation does occur, the temptation to think that it entails or necessarily engenders the formation of second-order desires derives perhaps from the fact that evaluation usually is accompanied by the formation of second-order desires. That this is so is hardly surprising. Why, after all, would we want to evaluate our desires if not because we care about what they are and hence could be moved to change them? But it cannot be the evaluation alone that would move us. As we've seen, what might move us is just what would move us to evaluate our desires in the first place. In other words, the second-order desires that we might be tempted to see as the products of evaluation are in fact derived from the desires that motivate the evaluation together with the evaluative beliefs.

6. Conclusion

Social scientists tend to fall into two camps on the question of normative motivation. In one camp are those (call them Sociologists) for whom it is not at all problematic that norms affect behavior – and in a way that does not involve instrumental rationality – and who take it for granted that a person with a normative belief, a belief that she ought to ø, must have some motivation to ø. In the other camp are those (call them Economists) who do not at all take these things for granted. But, ignoring the possibility that there is some distinctive form of moti-

vation at work here or some effect of norms distinct from the effects of sanctions, they try to assimilate all behavior to the economists' rational-choice framework, including behavior conforming to norms or consistent with normative beliefs that at first blush seems not to be the behavior of a rational egoist. Recently, a few political scientists sympathetic to rational choice theory have begun to take norms seriously, and a number of people from various disciplines have made large, interesting, but as yet embryonic and imprecise claims about 'social capital', of which norms are apparently one form (see Coleman 1990, ch. 12; Putnam 1992). Yet none of these people has tried to isolate the distinctive effects of norms on behavior – distinct, that is, from any extrinsic sanctions that back them up – or has seen that what I am calling normative motivation, or motivation by normative belief, is at all problematic.

Getting clear about these things seems to me to be a necessary preliminary to progress on a range of questions – including the following. Exactly what does 'social capital' consist in and just how does it facilitate the actions of individuals who collectively possess it? Do certain norms, in particular, facilitate mutually beneficial exchanges and, if so, how? (James Coleman says that "the relations of authority and of trust and ... norms ... are forms of social capital" but also that "social capital inheres in the structure of relations between persons and among persons" (Coleman 1990, 300-2). Is there more to social capital than social relations with the characteristics of a community, and does it have any effect distinct from the second- and third-party controls characteristic of community?) Under what conditions, and for what sorts of agents, is normative motivation or some other form of self-control likely to be effective? In what conditions or circumstances does it make sense to try to foster or strengthen the foundations for such forms of control?

What I have done here is merely preliminary to all this. I have insisted that we try to isolate moral motivation as such and that motivation by moral belief should not be taken for granted. And I have warned that in looking for motivational powers or effects of normative belief we must be careful to rule out the possibility that the motivation is one that is derived from other existing desires. Nor do deliberation and evaluation produce desires purely out of beliefs, normative or other, but rather involve the straightforward derivation of desires from beliefs and (pre)existing desires.

Having rejected internalism – the thesis that there is a necessary or conceptual or internal link between an agent's moral belief and her motivational state – and rejecting also the thought that there is no link at all between these two, we are left with the task of sorting out the contingent connections between moral belief and motivation. One obvious set of contingent links involves what have been called the emotions of self-assessment, notably guilt and shame (Taylor, G. 1985).[12] I

[12] It is odd of Brink (1989, 49) to offer *sympathy* as his only example of such a contingent link between having a moral belief and having some desire to comply with what one perceives to be one's obligations. Would not sympathy normally motivate independently of

could not here do justice even to these links, but I want in conclusion to suggest that guilt and shame, which are so often conflated, work in quite different ways, and shame plays one role in influencing our behavior that cannot be captured by the model of instrumental rationality. This is because, while guilt is something we experience when we have done something forbidden or have failed to do something required of us and is focused on a particular act, shame is experienced when we see our selves in a certain light, and experiencing shame therefore presupposes that we have selves, in a sense that I believe entails that not all our behavior can be instrumentally rational. Or so I would argue. That, however, is a long and complex story that I must tell elsewhere.

Bibliography

Benn, Stanley (1979), The Problematic Rationality of Political Participation, in: Peter Laslett/James Fishkin (eds.), *Philosophy, Politics and Society, 5th series*, Oxford

de Bretton Platts, Mark (1991), *Moral Realities*, London

Brink, David O. (1989), *Moral Realism and the Foundations of Ethics*, Cambridge

Cohen, Jean L. (1985), Strategy or Identity: New Theoretical Paradigms and Contemporary Social Movements, in: *Social Research 52*, 663-716

Coleman, James S. (1990), *Foundations of Social Theory*, Cambridge/MA-London

Copp, David (1991), Moral Realism: Facts and Norms, in: *Ethics 101*, 610-624

Etzioni, Amitai (1988), *The Moral Dimension*, New York

Foot, Philippa (1978), *Virtues and Vices*, Los Angeles

Frankfurt, Harry (1971), Freedom of the Will and the Concept of a Person, in: *Journal of Philosophy 68*, 5-20

Harrison, Jonathan (1976), *Hume's Moral Epistemology*, Oxford

Hollis, Martin (1977), *Models of Man*, Cambridge

Hume, David (1888), *A Treatise of Human Nature*, ed. Selby-Bigge, Oxford

Kreps, David M. (1990a), *Game Theory and Economic Modelling*, Oxford

– (1990b), Corporate Culture and Economic Theory, in: James A. Alt/Kenneth A. Shepsle (eds.), *Perspectives on Positive Political Economy*, Cambridge, ch. 4

McDowell, John (1978), Are Moral Requirements Hypothetical Imperatives?, in: *Proceedings of the Aristotelian Society, supp. vol. 52*, 13-29

Mackie, John L. (1977), *Ethics*, Harmondsworth

Milligan, David (1980), *Reasoning and the Explanation of Actions*, Brighton/Sussex

Nagel, Thomas (1970), *The Possibility of Altruism*, Oxford

Pettit, Philip (1987), Humeans, Anti-Humeans, and Motivation, in: *Mind 96*, 530-533

– /Michael Smith (1990), Backgrounding Desire, in: *The Philosophical Review 99*, 565-592

Putnam, Robert D. (1992), *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton

---

moral obligation? Surely, sympathy is not produced by a sense of moral obligation and is not what would usually provide the bridge from moral belief to motivation.

Ramsey, Frank (1978), *Foundations: Essays in Philosophy, Logic, Mathematics and Economics*, ed. D.H. Mellor, London

Smith, Michael (1987), The Humean Theory of Motivation, in: *Mind 96*, 36-61

– (1989), Dispositional Theories of Value, in: *Proceedings of the Aristotelian Society, supp. vol. 63*, 89-111

Snare, Francis (1991), *Morals, Motivation and Convention*, Cambridge

Sturgeon, Nicholas (1986), What Differences Does It Make Whether Moral Realism is True?, in: *The Southern Journal of Philosophy, supplement 24* (= *Moral Realism. Proceedings of the 1985 Spindel Conference*, ed. N. Gillespie)

Taylor, Charles (1985), What is Human Agency?, in: *Human Agency and Language: Philosophical Papers I*, Cambridge

– (1989), *Sources of the Self: The Making of the Modern Identity*, Cambridge/MA

Taylor, Gabriele (1985), *Pride, Shame, and Guilt*, Oxford

Taylor, Michael (1982), *Community, Anarchy and Liberty*, Cambridge