

PHILOSOPHICAL EGOISM: ITS NATURE AND LIMITATIONS

HANS BERNHARD SCHMID

Universities of Basel and St. Gallen

Egoism and altruism are unequal contenders in the explanation of human behaviour. While egoism tends to be viewed as natural and unproblematic, altruism has always been treated with suspicion, and it has often been argued that apparent cases of altruistic behaviour might really just be some special form of egoism. The reason for this is that egoism fits into our usual theoretical views of human behaviour in a way that altruism does not. This is true on the biological level, where an evolutionary account seems to favour egoism, as well as on the psychological level, where an account of self-interested motivation is deeply rooted in folk psychology and in the economic model of human behaviour. While altruism has started to receive increasing support in both biological and psychological debates over the last decades, this paper focuses on yet another level, where egoism is still widely taken for granted. *Philosophical egoism* (Martin Hollis' term) is the view that, on the ultimate level of intentional explanation, all action is motivated by one of the agent's desires. This view is supported by the standard notion that for a complex of behaviour to be an action, there has to be a way to account for that behaviour in terms of the agent's own pro-attitudes. *Psychological altruists*, it is claimed, are *philosophical* egoists in that they are motivated by desires that have the other's benefit rather than the agent's own for its ultimate object (*other-directed desires*). This paper casts doubt on this thesis, arguing that empathetic agents act on other people's pro-attitudes in very much the same way as agents usually act on their own, and that while other-directed desires do play an important role in many cases of psychologically altruistic action, they are not necessary in explanations of some of the most basic and most pervasive types of human altruistic behaviour. The paper

Previous versions of this paper were presented at the *Workshop on Altruism* at the University of St. Gallen on 31 May 2007, and at the universities of Bern and Hannover. The author wishes to thank these audiences as well as the editors and the anonymous referees for *Economics and Philosophy* for sharp comments.

concludes with the claim that philosophical egoism is really a cultural value rather than a conceptual feature of action.

1. EGOISM: BIOLOGICAL, PSYCHOLOGICAL, AND PHILOSOPHICAL

In English, just as in most other European languages, the word 'egoism' had already been in use for quite some time before Auguste Comte introduced its antonym into western thought. Ever since its invention, the term 'altruism' has had a hard time finding a place in explanations of human behaviour. There is a sense in which egoism seems natural and unproblematic in a way that altruism does not. Even Comte himself, one of the most ardent advocates of his terminological offspring, had to admit that egoism is genetically secondary, motivationally more stable and structurally less demanding than altruism (e.g. Comte 1851: 693f). A great number of thinkers have gone further in claiming that apparent cases of altruistic behaviour are really just some more sophisticated version of egoism. The reason for this is that given our theories about human behaviour, egoism has an obvious explanatory advantage over altruism. This is true both on the biological and the psychological level of the debate. In an evolutionary account of the biological world, any 'self' (be it a group, an individual, or a gene) is selected for its capacity to maximize its own reproductive fitness. This makes biological altruism seem a hopelessly self-defeating strategy: maximizing the reproductive fitness of *others* at the cost of their own, biologically altruistic selves would soon disappear from the scene of the Darwinian struggle for survival. On the psychological level, where the focus is on *motivation* rather than on fitness effects, a similar picture emerges. Psychological egoism is the claim that people act only in their perceived self-interest, a notion that receives strong support from our standard view of human motivation, according to which people basically seek to maximize their own expected utility.¹ This makes psychological altruism seem highly problematic: if an agent acts in the interest of another person at costs to herself, the question arises of why she did it, and it is often claimed that she *must* have expected to get something out of the deal for herself after all, be it in the form of the infamous 'warm glow' (Andreoni 1990), some sympathetic satisfaction (Becker 1986) or simply in avoiding the negative arousal (e.g. pricks of conscience) she expects to experience were she to act differently.

¹ Psychological egoism is to be distinguished from biological egoism because people's self-interests may differ considerably from their reproductive fitness. To use Ernst Fehr's example (Fehr and Fischbacher 2003): smoking a cigarette is psychologically egoistic in that it (myopically) optimizes the agent's well-being, while it is biologically altruistic since it decreases the agent's reproductive fitness and contributes to that of other agents by eliminating one competitor for resources from the scene.

In spite of the explanatory advantage of egoism, altruism has started to receive increasing support on both the biological and psychological level in the last few decades. On the biological level, kin selection theories (Trivers 1971) have shown how the altruistic behaviour of one individual towards her kin might increase the reproductive fitness of her genes; a complex of behaviour that is altruistic from the point of view of the individual appears as compatible with the 'selfish gene'-view of evolution. Going one step further, group selection theory (Sober and Wilson 1998) argues that groups may be seen as separate units of selection, and that under some (very specific) circumstances altruistic behaviour even among non-kin is evolutionarily stable in that it increases the group's fitness. On the psychological level, Dan Batson has gone at great length showing that 'pure altruism' (i.e. actions whose ultimate goal is the benefit of others) is empirically plausible (e.g. Batson 1991); Eliot Sober and David Sloan Wilson (1998) have argued forcefully that while the existence of psychological altruism might turn out to be impossible to prove, it is an evolutionarily plausible hypothesis.²

In all of these interpretations, as well as in Philip Kitcher's important work on the topic (e.g. Kitcher 1998), psychological altruism appears as a special mode of motivation. Psychological altruists do not seek to improve their own well-being, however broadly conceived, but act on desires whose ultimate goal is to promote other people's well-being, independently of whether or not they get anything out of the deal for themselves. In other words, the agent herself does not figure in the content of the desire by which the psychologically altruistic action is motivated. Using Philip Kitcher's term, one may call these *other-directed desires*.

This paper addresses a further level of the debate on which egoism is still widely taken for granted. I am concerned with a sense in which even psychologically altruistic action, insofar as it is motivated by other-directed desires, might still be called egoistic in that it is *her own* desire on which the psychological altruist acts, and not her beneficiary's. Even though the promotion of the beneficiary's interests, desires, or well-being rather than her own will figure in the *content* of the desire in question, the psychological altruist promotes her beneficiary's interests only insofar (and to the degree to which) she herself wishes to do so. However non-selfish her interests may be, the motivational agenda behind her behaviour is still her own. In order to direct his reader's attention to this notoriously elusive feature and to distinguish it from psychological

² In Sober's and Wilson's example: a parent with the capacity to care for her children, even in the hopefully rare cases in which she happens not to feel like it at all, might be more efficient than the one who is only motivated by her parental inclinations; psychological altruism appears as a back-up system for the case of the breakdown of the regular psychologically egoistic operating mode.

egoism, Martin Hollis introduces the label 'philosophical egoism', which seems appropriate because the issue at stake here is not a question concerning the *kind* of motivation, but rather a question of its conceptual structure, which falls firmly into the domain of philosophy:

Even if there is no psychological egoism, there is still prone to be a philosophical egoism in all accounts of what moves human agents. It surfaces when we ask how exactly a preference for *x* over *y*, or a calculation that *x* offers more utility than *y*, moves someone to act. In so far as preferences are a newer name for what used to be called passions, the classic answer is that the agent expects to gain greater psychological satisfaction. Since not all sources of satisfaction are self-centered, there is room for many desires and many ways to satisfy them. If 'self-interest' is construed in this broader sense, we can still hold that 'every agent is actuated solely by self-interest'. But what is then meant is that all action comes about as the stock desire/belief model suggests, by the prompting of desire, tempered by the agent's beliefs about alternative ways to satisfy it. Crucially, Adam is moved solely by what Adam wants and Eve solely by what Eve wants. Call this philosophical egoism. (Hollis 1998: 20f)

Aside from the label, the distinction between psychological and philosophical egoism as such is not Hollis' invention. Already in a paper from 1958, Joel Feinberg had shown that no convincing conception of 'desire' allows us to derive psychological egoism from the assumption that actions are motivated in the agent's own desires (Feinberg 1958/1995). However, since Feinberg's aim was limited to defending psychological altruism, he used this argument only to prove that it is wrong to stick to psychological egoism from fear of having to reject philosophical egoism, as often seems to occur, and that there is a way of conceiving of psychological altruism which is compatible with philosophical egoism. In his paper, Feinberg does not question philosophical egoism. He simply points out that philosophical egoists are psychological altruists to the degree that their desires are other-directed, thereby affirming the limited conception of psychological altruism still accepted in the current debate.

In the following, I will not take issue with Feinberg's fundamental insight, which I take for granted. The aim of this paper is to challenge philosophical egoism *on its own terms*, i.e. as a kind of egoism that must be distinguished from psychological egoism, and this will allow me to sidestep the confusion that Feinberg has already so thoroughly cleared up. I shall challenge philosophical egoism as a general theory about the structure of human action, and argue for the possibility of philosophical altruism, i.e. action which is not motivated by the altruist's own other-directed desires, but by the volitive or conative attitudes of others. This is no easy task, as the theory behind philosophical egoism seems to be even more formidable than the theories behind biological

and psychological egoism. On the philosophical level, the objection to be met is not an empirical claim about the development of the biological world, or some folk-psychological assumption about human motivation but rather a purely *philosophical, conceptual point* about the very nature of action. Biological altruism may seem *highly improbable* in the light of evolution theory, and psychological altruism appears as rather *implausible* in the light of the standard view of human motivation; the idea of philosophical altruism, however, is faced with the objection that it is simply an *inconsistent* idea and, as a matter of pure conceptual necessity, *impossible*. The reason is this: it seems plausible to say that for a complex of behaviour to be an action, there has to be a description under which the agent *wanted* to do it; according to mainstream action theory, actions are *identified* by the *pro-attitudes* (Donald Davidson's term) of the agent for whom they act as a motivation. This makes philosophical egoism appear as a structural feature of *any* action, however altruistic it may be. The basic argument that I shall develop in this paper in order to meet this objection is that this view is not so much mistaken as it is imprecise. My argument relies on the distinction between *intentions* and *desires*, the two of which are usually lumped together under the Davidsonian label 'pro-attitude'. While an *intention* needs to be the agent's own, the motivating *desire* does not, or so I shall argue. This leaves ample space for philosophical altruism: philosophical altruists act *intentionally* on other people's desires or intentions, but the reason for their doing so is not to be found in some other-directed desire, but rather in the other's volitive or conative states of which altruists are empathetically aware. Empathy plays the exact same structural role in philosophically altruistic action as the agent's awareness of her own desires does in philosophically egoistic cases, but extends the class of possible motivating reasons for action beyond her own 'subjective motivational set'.

I shall proceed as follows. First, I will dispel the idea that philosophical altruism is simply a confused notion that appears reasonable only to philosophically untrained minds (as it is sometimes claimed in the received literature) by introducing three philosophers who have argued for the possibility of philosophical altruism. Looking at their views will also help us to get a closer grip on why almost all philosophers (including some of those discussed) ultimately shy away from this notion and resort to other-directed desires explanations ('the paradox of philosophical altruism'). In the next section, I shall try to establish the fact that, in spite of these conceptual worries, there is some *intuitive* plausibility to the idea of philosophical altruism. For this purpose, I shall suggest a fundamental *shift of focus* in the debate. The paradigm cases of altruistic behaviour discussed in the received literature include examples such as donating to charities, acting as a Good Samaritan, or sacrificing one's life for others. I propose to shift away from such heroism and consider instead

spontaneous small-scale, low-cost, cooperative everyday behaviour, such as holding the door open for other people or moving aside to make room for another person on a park bench, both of which might seem merely to be routine acts of politeness rather than cases of proper altruism. I shall argue that many such acts are genuinely altruistic rather than norm-guided routines, and that in many of these cases, philosophical altruism seems intuitively more plausible than other-directed desires explanations. In Section 4, I will turn from articulating intuitions to revisiting the conceptual problem encountered in Section 2. I will argue that the paradox can be resolved, and that philosophical altruism is compatible with our standard conception of action, once it is understood correctly. My argument relies on the distinction between what I propose to call 'intentional autonomy' and 'motivational autarky'. Section 5 analyses the role of empathy and interpersonal identification. The concluding Section 6 addresses the question of the true nature of philosophical egoism. My claim will be that philosophical egoism is really a deep-seated *cultural ideal* rather than a conceptual feature of action. Acting exclusively on one's own motivating desires is part and parcel of our idea of a fully developed and self-dependent person, and this in turn is compatible with the fact that, very often, actual agents do not conform to this ideal.

2. THE PARADOX OF PHILOSOPHICAL ALTRUISM

Philosophical altruism is rarely taken seriously in the current literature. In those few cases in which the issue comes up, it is usually treated as a mere conceptual scam or the result of philosophical confusion. Thus Sober and Wilson (1998: 223) argue that it is simply a mistake to define egoism in terms of 'being motivated by one's own desires', and that this results in a 'spurious' and 'short-circuited' view of altruism. The undertones of Philip Kitcher's remarks on the topic seem even harsher. Kitcher appears to think that only non-philosophers could be so naïve as to think that there is more to the problem than mere conceptual confusion; he calls the idea of philosophical altruism a 'mistake' which, in a somewhat opaque dialectical move, he deems 'illuminating' because it 'distorts a genuine insight' (Kitcher 1998: 291). The genuine insight at stake is basically Feinberg's (1958/1995): it is that not all desires are self-directed. For Kitcher, just as for Sober and Wilson, it is clear that altruists, just like any other agents, are motivated by their own desires, although their desires are other-directed rather than selfish. According to these authors, just as for many others, the question of egoism and altruism is not a question of the 'owner' of the motivating desire, but rather a question of whether the agent himself or another person figures in its content. Yet the notion of philosophical altruism – if not the term – is neither new nor simply a

product of confusions occurring in philosophically untrained minds only. There have been some philosophers who have, to some degree at least, argued for this concept, and looking at some of their views might be a good point to start. I have selected three examples.

Arthur Schopenhauer may seem a problematic example, as he endorsed a narrow conception of motivation with hedonistic underpinnings that exclude the kind of psychological altruism at the centre of the current debate. Thus one might suspect Schopenhauer's endorsement of philosophical altruism to be a result of the fallacy identified by Feinberg. However, even if Schopenhauer was mistaken in excluding 'classical' psychological altruism, it seems wrong to presume that he might not have been onto something important in his account of philosophical altruism. Here is the crux of his argument in *On the Basis of Morality* (1840/1995): The only motive of the will, Schopenhauer claims, is either pleasure or suffering. Action is egoistic to the degree that the agent's will is moved by her own pleasure or suffering. Egoistic action is either morally neutral or unethical. Moral action requires altruism (though the term is not used by Schopenhauer). Action is altruistic to the degree that the beneficiary's pleasure or suffering is the altruist's *immediate* motive in the exact same way her will is moved by her own pleasure and suffering in all other actions. Thus the basic problem for an account of altruistic action, in Schopenhauer's view, is to show how another person's psychological states can *directly* motivate the altruist's action without any extra motive of hers interfering in the process. Schopenhauer does claim that this is in fact possible, and that compassion provides the answer to this question. But he also freely admits that 'this process is most puzzling, and indeed mysterious', as it blurs the distinction between persons (Schopenhauer 1840/1995, §16).

A second example is to be found in Thomas Nagel's *Possibility of Altruism* (1970), where Nagel claims that 'an appeal to our interests, or sentiments, to account for altruism, is superfluous. (...) There is, in other words, such a thing as pure altruism (though it may never occur in isolation from all other motives). It is the direct influence of one person's interest on the actions of another' (1970: 80). Nagel does not speak of *desires*, but rather of *interests*; but it is clear from the context that he is concerned here with *motivational* states. This is clear from the following passage, in which he anticipates a worry his critics may have concerning his previous claim: 'since it is I who am acting, even when I act in the interest of another, it must be an interest of mine which provides the impulse. If so, any convincing justification of apparently altruistic behaviour must appeal to what *I* want'. Nagel does not grant this objection. But, as he adopts a Kantian view of practical reason, he also does not provide a straightforward answer as to how other people's interests may prompt an altruist's action *directly*, and he even follows Kant

in his rebuke of compassion-based accounts of altruism. Yet still, the claim stands: as far as the motivational input for altruistic actions is concerned, the altruist's own psychology may be the wrong place to look; any appeal to an altruist's own motivational agenda might simply be superfluous, and the other person's interests might just be enough.

My last example is Amartya K. Sen's notion of 'committed action'. As early as his 'Rational fools' (1977), Sen argues explicitly against the view that committed actions can be accommodated within a preference-based framework simply by widening the scope of the agent's preferences. Committed action, he claims, involves 'counter-preferential choice', suggesting a behaviour that cannot be explained by the agent's own preferences, however widely they are conceived. In 'Goals, commitment, and identity' (1985/2002), Sen casts this claim in terms of goals rather than preferences; however, as goals can be seen as the conditions of satisfaction of *desires*, his considerations are directly pertinent to the question at issue here. Sen argues in this paper that it is a mistake to assume that 'a person's choices must be based on the pursuit of her own goals'. Committed agents, he suggests, may act directly on other people's goals, without making them their own. Sen points out that one person's *identifying* herself with another might play a role here, but he, too, clearly articulates the worries he expects his critics to have: 'It might appear that if I were to pursue anything other than what I see as my own goals, then I am suffering from an illusion; these other things *are* my goals, contrary to what I might believe' (1985/2002: 212).

Thus even a cursory look into the literature reveals that contrary to what Sober, Wilson and Kitcher seem to think, the idea of philosophical altruism has crossed many philosophically acute minds.³ But it is equally clear that neither Nagel nor Sen offers a straightforward conception of philosophically altruistic action, limiting themselves instead to the view that there is something wrong with philosophical egoism. Schopenhauer, by contrast, does elaborate on his view in some of his other writings, but since his ultimate metaphysical conclusion is that the difference between persons is only a matter of *appearance* and that 'in ourselves' we are really *one and the same* (cf. Schopenhauer 1849: 625), such an elaboration may not lend his notion of non-selfish behaviour additional plausibility – at least as an account of *altruistic* action (to the same degree that we are really one at some deeper metaphysical level, all action, be it motivated by one's own desires or by another's, is ultimately selfish).

Clearly, the problem with the notion of philosophical altruism is not empirical, but conceptual. In the chapter on Egoism and Altruism in his *Introduction to the Sciences of Ethics* (1892), Georg Simmel gives one

³ Another clear and well-argued example is Paprzycka (2002).

of the clearest (if somewhat idealistic) statements of why the idea of philosophical altruism might be mistaken a priori:

Just as all objects of possible consideration are only in my imagination, since I cannot outrun my ego in my thoughts, I could never do it in practice either. All imagining is *my* imagining, and likewise, all willing is *my* willing, and I could not possibly pursue anything but my own goals. Just as, according to the Kantian conception, the things in themselves do not enter my mind, the interests of other people cannot determine my will in action. Real objects exist for me only if they become subjective and thus present in my imagination. In the same way, other people and their interest are relevant to me only when mediated through my own interests. Only by making another person's interests my own can my will acquire any altruistic content. (Simmel 1892, Vol. 1, Ch. 2)

In the terminology of present-day action theory, Simmel's intuition can be cast more sharply and without idealistic overtones. One basic role of motivational states is that they *rationalize* action; they are the *reasons* that distinguish actions from other kinds of events that have only *causes*. By identifying actions, reasons for action (which split into beliefs and desires) also identify the *agent*. In Donald Davidson's words: 'R is a primary reason why an agent performed the action A under the description d only if R consists of a pro attitude *of the agent* toward actions with a certain property, and a belief of the agent that A, under the description d, has that property' (Davidson 1963: 687; my emphasis). Thus it seems that philosophically altruistic action is a simple *contradiction in terms*. If the altruist is to be the agent of her own behaviour, the primary reasons for that behaviour have to be *hers*. Thus her behaviour cannot be philosophically altruistic. (Remember that *ex hypothesi* such behaviour is not to be rationalized by the altruist's own pro-attitudes, but rather by the beneficiary's; therefore, the altruist's behaviour would not instantiate her own actions, but rather the beneficiary's.) An altruist's behaviour can be either her own action, or it can be philosophically altruistic, but it cannot be both. Since it is plausible to assume that an altruist's behaviour *does* instantiate her own actions (the metaphor 'lending a hand' should not be considered more than just that: a metaphor), it follows that there is no philosophical altruism. Philosophers like Schopenhauer, Nagel and Sen were simply on the wrong track in the passages quoted above. There might be psychological altruism, in terms of actions based on other-directed desires, but philosophically, we're all really egoists – or so it seems.

3. EVERYDAY ALTRUISM

Having addressed the conceptual problem with philosophical altruism, I will now try to show that, in spite of these philosophical worries, there

is a great deal of plausibility to philosophical altruism at the *intuitive* level. In order to do so, I recommend a shift of focus concerning the kind of phenomena taken into consideration. In the received literature on altruism, the paradigm cases are donating to charities, helping Jews in Nazi Germany, acting as a Good Samaritan, or the famous WWI Lieutenant throwing himself onto the grenade that has fallen into his trench in order to protect his comrades. By contrast to such heroism, the kind of behaviour analysed in this paper is of a much less spectacular kind. As our paradigm case, I choose the following example. In a session of the Economic Science Association at the ASSA-meeting in Chicago early in 2007, the economist and behavioural scientist Herbert Gintis opened his talk on altruism with a simple case of everyday behaviour that he had just witnessed. Standing with his suitcases before closed doors in front of the conference building and unable to find the open-door button, some passer-by who observed the scene had taken it upon herself to press the button for him, leaving the scene immediately after having helped without even waiting to be thanked. Perhaps Gintis is right and more attention should be devoted to behaviour of this kind in the debate on altruism. Such behaviour is pervasive in social life; it certainly does occur in intimate relationships, too, but its special status becomes even more visible in the anonymity of the public domain: people holding doors open for strangers carrying suitcases, passengers helping each other to lift baby carriages into and out of trains, people moving aside on their benches so that other people can sit down too; commuters on railway platforms facilitating other people's passage by moving out of their way; passengers assisting each other lifting their suitcases to and from carry-on luggage trays, people picking up objects for other people.

Such behaviour is considerably different from the kinds of examples usually encountered in the literature. At least three distinctive features are immediately apparent. First, it is essential to the paradigmatic cases of altruism found in the received literature that the altruists incur some cost (be it time, effort, money, or, in the extreme case, one's own life). Some degree of *self-sacrifice* is usually taken to be essential for an action to be altruistic. By contrast, our examples seem to be marked by *indifference*. It is true that, in actual fact, the benefactors do incur some costs, but they are minimal and they seem to play no role in the benefactor's own perception of the situation. Where the stakes are high, such behaviour usually disappears; it might be difficult to find a person ready to hold a door open for another passenger when she knows that she may miss her train as a result. Such behaviour occurs in low-cost situations only, or so it seems. Second, such acts seem to be, to a large degree, *non-premeditated*. These benefactors act more or less *spontaneously* and perhaps even *unthinkingly*, following well-established routines in their everyday lives. This is very different from cases such as the donor's, where some

conscious deliberative process of weighing one's own interests against the beneficiary's seems to be essential. Third, there is a fundamental difference in the kind of attitude at work between the benefactor and her beneficiary; the view of other underlying such behaviour is *superficial*. Classical acts of altruism are marked by some sort of *care* or *concern* for the beneficiaries. This entails that the benefactor has some conception of the beneficiary's needs, which in vicarious or patronizing forms of altruistic action might differ from the beneficiary's own, as well as from his manifest desires and intentions. As opposed to this, the behaviour of the above agents is not guided by an understanding of any of the beneficiary's deeper needs, but rests entirely at the level of their immediate and manifest *goals*. These benefactors support their beneficiaries in whatever they seem to be trying to do, and this does not involve any further evaluation of these goals, which seems to make these cases a matter of manner rather than of morals. In short, the phenomenon is this: other-directed, spontaneous, routine-like and apparently non-deliberative action in which other people are supported in the pursuit of their immediate goals in low-cost situations. In what follows, I shall call such behaviour *everyday altruism*.

Looking at these differences, one might doubt whether or not such behaviour should be taken as cases of altruism. Especially philosophers working on ethics tend to have rather high expectations for altruistic behaviour, demanding some sort of *concern* addressing the deeper needs of the other rather than just a tendency to spontaneous cooperation, and some degree of self-sacrifice rather than just minimal cost assistance. Be that as it may, it seems clear that such behaviour does nicely fit the phenomena Auguste Comte had in mind when he coined the term, and experimental economists, who have now started to claim the label for themselves, will have no difficulty accepting this classification (cf. Fehr and Fischbacher 2003). After all, the behaviour in question does benefit another person, and it does come at a cost to the benefactor, however minimal it might be. The question is: why do people behave this way, especially where the type and the anonymity of the situation seems to exclude reputation effects and sanctioning? The standard account of human motivation recommends looking for psychological rewards or costs, and indeed the effects of grateful smiles should not be underestimated; but in many cases (such as in Gintis'), everyday altruists do not even wait around to be thanked. As far as psychological costs are concerned, it is certainly true that we are creatures with a tremendous capacity for internal negative sanctioning (imagine the pang of shame you feel when you realize that you've been observed picking your nose even by a complete stranger), but as far as everyday altruism is concerned, no such taboo seems to be involved: sometimes people do it, very often they do not, and neither warm glow nor pangs of shame seem to account for the difference.

As far as the question of motivation is concerned, it is usually good advice to ask the agents. This is not to say that agents are always truthful concerning their motivation, and especially if one is partial to psychoanalysis, one might even allow for cases in which the agents are simply *incompetent* concerning the question of their own ultimate motives. Also, there might be a difference in terminology; philosophers sometimes use terms such as 'desire' simply for behavioural dispositions rather than for some internal psychological entity. But as far as normal, non-pathological cases of action and standard usages of motivational vocabulary are concerned, it seems that the agents themselves are in a privileged position. So it might be worth thinking about the kinds of answers everyday altruists might come up with when asked about their motivation.

As far as standard cases of altruistic actions are concerned, such research has already been carried out by social psychologists. When 'classical' altruists who have donated to charity or done volunteer work were asked why they did so, they usually answered that they 'wanted to do something useful' or that they 'wanted to do good deeds for others', or something along these lines (Reddy 1980 quoted in Sober and Wilson 1998: 252). Such self-reports are, of course, in perfect tune with classical accounts of psychological altruism: the ultimate goals that these people cite are other-directed, as the agents themselves do not figure in the content of their motivation. On the philosophical level, such motivations are clearly egoistic: the desires cited by these altruists are *their own* desires. What motivated their action was what *they* wanted, which corresponds to the view of philosophical egoism that the only motivational base for action is self-interest, if self-interest is taken in the purely formal sense of ownership rather than content, i.e. in the sense that the interest at stake is the agent's own rather than anybody else's (remember Martin Hollis' definition of philosophical egoism in the first section above). To adherents of standard action theory, this result will come as no surprise, because for them this is simply a matter of conceptual necessity and so not up for empirical falsification. However, there seems to be a way in which cases of *everyday altruism* can be explained in ordinary language that does not fit so well with philosophical egoism. I do not know if any such work has been carried out in social psychology, so I will have to rely on intuitions about ordinary language which I can only hope the reader also shares.

Imagine asking Herbert Gintis' helper why she pushed the open-door button for him. It is quite possible, of course, that she would say that she wanted to render that man a service, or that she simply wanted to be polite, or that she simply couldn't bear the sight of the man's helplessness, thereby citing some other-directed desire of her own. But there is something slightly artificial about such explanations. It seems much more plausible that her answer to the question would simply be:

'Because he [Gintis] couldn't find it.' Similarly, if one asked a person on the park bench for what reason she had moved aside when another person approached the bench, she would probably say 'because that person wanted to sit down, too' rather than 'because I wanted to make space for him to sit down beside me', or 'because I wanted to be nice to him' or something of that sort. The decisive difference is this: in explaining the behaviour in question, these reports cite *other people's pro-attitudes* rather than the agent's own. As opposed to donors, volunteer workers, or other classical altruists, everyday altruists seem more likely to explain their behaviour in terms of what *other* people want rather than in terms of their own desires. Insofar as this is true, the possibility arises that philosophical altruism might not after all be nothing more than a confused philosophical idea in the minds of such authors as Arthur Schopenhauer, Thomas Nagel and Amartya Sen; it may also be part of the everyday altruist's own self-understanding, thus adding further weight to the idea.

However, even if this intuition concerning ordinary linguistic practices is accepted as plausible, there are still alternative interpretations to consider. One way to make such manners of speaking compatible with philosophical egoism relies on the difference between *motivation* and *justification*. When they explain their behaviour in terms of the pro-attitudes of *other* people rather than their own, one might think that everyday altruists are pointing out those reasons in the light of which their actions are *justified* rather than saying anything about the motivating reasons for those actions. The distinction between justifying and motivating reasons (cf. Pettit and Smith 2004: 270) is *fundamental* insofar as agents might be *motivated* by reasons which they do not take to be *justified*, such as the case of the unwilling addict who acts on his desire to take the drug without taking the satisfaction of his desire to be a goal worthwhile pursuing. Such behaviour is *rationalized* by the motivating desire without being fully *rational* for lack of a justifying reason. In normal cases of action, however, justification and motivation do not come apart entirely. In the Kantian view, it is *because* she sees it as worth doing that a rational agent's will is moved to perform an action. In the Humean view, some further motivation is assumed, such as the desire to do the right thing. Thus the objection to the view that the above ordinary language examples express philosophical altruism is that these everyday altruists only refer to *justifying* reasons while remaining silent about their motivational structure, which they simply take for granted (everyday altruists do not deem it necessary to point out that they are motivated to do the right thing). The other's intention or desire did not *motivate* their helping behaviour; rather, it was the reason in the light of which they were *justified* in wanting to intervene.

In order to assess the strength of this alternative view, we need to alter the situation so as to make sure that the explanation given by an

everyday altruist is focused on motivation rather than justification. The following modification was suggested to me.⁴ Consider again Herbert Gintis standing in front of the closed door with his helper approaching the scene. But now suppose that there is another person, the helper's colleague, whom the helper knows to be familiar with the opening mechanism, and who is closer to the button than the helper herself. The helper sees that her colleague is aware of the fact that Gintis cannot find the button. But the colleague doesn't seem to bother, so the helper steps in and pushes the button herself. What would she say now, were she asked why she did so?

Before considering possible replies, a word on how this modification helps us to focus on motivation rather than justification is in order. According to the contrastive nature of any explanation (Garfinkel 1981: Ch. 1), the question 'why did you push the button', as asked of the helper, acquires a different meaning in these altered circumstances. Now, the question is not so much 'why did you push the button *rather than doing nothing*', but 'why did *you rather than your colleague* push the button?' This change in the background of the question moves the focus from justification to motivation, because as far as justification is concerned, both the helper and his colleague are in the same position: both had equal justifying reason to intervene. Therefore, pointing out the justifying reason would do nothing to explain the difference in their behaviour. Thus it seems that, in this situation, the helper's reply will finally be a clear indication of whether or not she sees herself as a philosophical egoist: the reasons she quotes will be her *motivating* reasons. If she sees herself as a philosophical egoist, her reply to the question would have to be something along the lines of 'I pressed the button because I wanted to help/wanted to be polite (while my colleague did not seem to have any such desire).' It does not seem, however, that such a reply would have to be given. It seems at least equally natural to expect an answer like 'because he [Gintis] wanted to enter the building and my colleague didn't bother to help him.' Again, this explanation does not cite the altruist's own pro-attitudes, but rather someone else's. As far as this is convincing, it seems that ordinary language and folk psychology do not unequivocally support philosophical egoism. It remains a remarkable fact about everyday life that, where motivation is concerned, people often explain their behaviour in terms of other people's pro-attitudes rather than in terms of their own, frustrating, to some degree at least, the attempt to make sense of their behaviour in terms of their own psychology. Thus it might be worthwhile taking a closer look at the paradox of philosophical altruism. Is there a way to fit philosophical altruism into a reasonable account of action,

⁴ This example is courtesy of an anonymous referee for *Economics & Philosophy*.

absolving such cases of charges of sloppy talk, mere self-deceptions, or false consciousness?

4. THE PARADOX RESOLVED

As noted above in Section 2, the difficulty is a conceptual one. For a complex of behaviour to be an action, there has to be a description under which the agent *wanted* to do it. This is to say, there has to be a way to make sense of the behaviour in question in terms of the agent's own pro-attitudes. Once more, Feinberg's fundamental insight should be remembered: the claim at stake here is neither that people are autistic and act in complete disregard of other people's wishes (people do take other people's wishes into account in the pursuit of their actions), nor is it that people act only in the pursuit of their own selfish goals (people may well be psychological altruists and act on nothing but their own desire to fulfil another person's wish, without wanting to get anything out of the deal for themselves). The egoism in question here is not psychological, but philosophical; philosophical egoism seems to be built into our very notion of action. Were a subject to act *directly* and *exclusively* on another person's pro-attitude, i.e. without having any volitional agenda of her own, her behaviour would be rationalizable only in terms of that other person's pro-attitude and would thus be this other person's action rather than the subject's own. One might call such a hypothetical subject an *intentional zombie*: her behaviour would instantiate entirely another person's agency; she would be behaving entirely on that other person's strings. Intentional zombie-ism often occurs in sci-fi novels and in the self-reports of schizophrenics. It is not, however, a feature of everyday life, and certainly not present in the cases of everyday altruism mentioned above. The behaviour of everyday altruists does instantiate their own actions. But how, then, could it possibly be philosophically altruistic? The solution I propose hinges on the distinction between intention and desire. There is a sense in which everyday altruists do what they *want* and *because* they want to, but this 'wanting' should be understood in conative rather than motivational terms.

Were one to ask Herbert Gintis' helper whether she had *wanted* to push the open door button (rather than acting, in a Manchurian-Candidate-like way, on Gintis' strings), her reply would surely be positive. But the term 'wanting' is notoriously ambiguous, oscillating between 'aiming at' and 'being motivated to'. Labels such as Davidson's 'pro-attitudes' or Bernard Williams' (1981) 'subjective motivational set' lump together a person's motivational states (such as inclinations, urges, and desires) and her practical commitments (intentions, plans, and projects). At this point, it is important to take a closer look at the relation between the volitive and the conative elements of an agent's 'subjective

motivational set'. How are desires and intentions related? The received literature distinguishes two types of relation. The first, constitutive account ties intentions closely to desires; desires are *constitutive* of intentions in that they are the volitional component by which an agent cannot intend to A without *wanting* to A.⁵ This constitutive reading of the relation between intention and desire, in which desire is really a conceptual component of intention, has to be distinguished from a motivational reading, in which desire and intention are related in rational and perhaps causal terms rather than in constitutive terms. In this motivational sense, desires are the *rational base* on which intentions are formed. The fact that a person is thirsty is the *reason* why she intends to have a drink. Here, the desire logically *precedes* the intention and provides the *motivating reason* for which an intention is formed. Thus there is a further ambiguity to be cleared up in the assumption that for a complex of behaviour to be an action, a linguistically competent agent has to be able to come up with a description under which she *wanted* to do what she did. This assumption is unproblematic insofar as it means that she has to be able to cite some 'constitutive desire', which really amounts to nothing more than pointing out the intention; it is not unproblematic at all, however, if it is taken to mean that she has to be able to cite some motivating desire of her own.

When Gintis' helper says – as she certainly would, were she asked – that she *wanted* to do what she did (after all, she was not forced to do so by Gintis' telekinetic powers), she clearly refers to a conative attitude. What she means is that she did what she did *on purpose*, i.e. *intentionally* (rather than behaving in a way beyond her control). This does not conflict with her claim that, as far as her *motivation* is concerned, the reason for her action is not to be found in her own 'subjective motivational set', but rather in Gintis'. The intention is hers, and this involves a constitutive desire. The motivational desire on which her intention is formed, however, is not hers.

Thus the claim that philosophical altruism is compatible with the idea that for a complex of behaviour to be an action it has to be possible to make sense of that behaviour in terms of the agent's own pro-attitudes rests on the distinction between two readings of this Davidsonian claim. The weaker reading, which I recommend, and which does not entail philosophical egoism, is what I propose to call *intentional autonomy*. Intentional autonomy requires that under normal circumstances (barring reflex behaviour and similar cases) an individual's behaviour instantiates

⁵ It should be noted that a constitutive reading of the relation between intention and desire requires a wide conception of desire. If desire is understood in the narrow sense of a mental state with a content the thought of which induces some positive affective reaction (Schueler 1995), it seems that nobody would ever intend to keep their annual dentist's appointment.

his or her own action. This excludes intentional zombie-ism. In order to endorse this claim, we need not, however, accept the stronger reading that is usually given to the Davidsonian principle, which amounts to philosophical egoism and which I claim to be false. I propose to label this reading *motivational autarky*. This reading claims that any motivational explanation of an action ultimately has to bottom out in the agent's own desires. According to this reading, agents may take into account other people's desires in whatever way they like, but they act on those desires only if and insofar as they have a desire of their own to do so, i.e. a desire which may be other-directed but is their own in the formal sense of Hollis' definition of self-interest. I call this reading motivational autarky because the image of agency it projects is somewhat similar to the view of closed economies. The idea is that the only motivational resources on which agents may draw are their own.

Before discussing this distinction between intentional autonomy and motivational autarky further, a word on how this opens up space for philosophically altruistic action is in order. If action conceptually requires only intentional autonomy, then there is nothing paradoxical about the notion of philosophically altruistic action. Philosophical altruists are agents in their own right, and not just something like the extended bodies of their beneficiaries, insofar as the intention on which they act is *theirs*. However, the motivational explanation of their action does not bottom out in any of their own wishes, but rather in their benefactor's. Philosophical altruists are intentionally autonomous, but motivationally non-autarkical. Philosophical altruists are agents whose intentions are formed by deliberative processes not limited to their own psychological states. Such agents sometimes treat other people's desires in the exact same way they do their own, considering them potential reasons to form an intention.

5. EMPATHY AND IDENTIFICATION

This solution to the paradox of philosophical altruism raises new questions. How can another person's desire or intention become the reason for a philosophical altruist's intention if she has no conforming desire of her own? Part of what makes this process so 'mysterious' (Schopenhauer's word) lies in how desires are usually conceived of. Some philosophers take desires to be mere behavioural dispositions. This has the disadvantage of making it difficult to accommodate cases in which motivation and action seem to come apart (where agents fail to act on what they themselves take to be their strongest desire, a phenomenon that seems to be rather widespread in everyday life). The main alternative is to conceive of desires in phenomenological terms; not all desires need to be conscious, but in order for a mental state to count as a desire, it has

in principle to be accessible to consciousness. If a desire is conscious, it must involve some, however vague, awareness (or 'representation') of the desired object or state of affairs, and it is felt as a push or pull towards that object or state of affairs. Conceiving of desires in phenomenal rather than dispositional terms makes philosophical egoism plausible in a way philosophical altruism is not. In the first case, it seems clear how the desire's motivational push brings the agent to form an intention: it is he who *has* the desire after all. In the case of the philosophical altruist, things seem different: the motivational push is an event in another person's psychology which is not even directly observable and accessible to the conscious experience of another person. The motivating desire and the action-guiding intention are events in different *monads*, to use Edmund Husserl's term, making it entirely unclear how the first event could ever motivate the second. How could anyone ever be *directly* moved by a desire she or he does not have?

It has been pointed out repeatedly in the history of philosophy that there is something deeply wrong about this whole way of conceiving of practical reason, and a closer look quickly reveals that the issue is not only the way in which this makes philosophical altruism implausible, but philosophical egoism as well. Kantians have never ceased pointing out that the idea that our own desires enter our deliberative processes as reasons is anything but unproblematic; in fact, for them it is doubtful whether *any* of one's desires could ever be, in itself, a reason for action. How could a desire acquire the status of a reason for an agent? Citing one's desire to have one's desires fulfilled does not help because it sets off a potentially infinite regress, and it is at odds with Harry Frankfurt's (1971) observation that, in many cases, we act intentionally on desires which are *in conflict* with second-order desires. Be that as it may, it should be remarked that the question of how our own desires move us to form intentions might not be quite as unproblematic as the received view has it.

Conversely, the fact that other people's pro-attitudes may function as ultimate motivating reasons in a person's deliberative processes might not be quite as mysterious as it might seem. In the received literature – most famously in Husserl's phenomenology – the relatively recent term *empathy* has been used to point out how this may come about. As far as we empathize with other people, we are *aware* of and *affected* by their pro-attitudes. At this point, it might be useful to remind the reader of the fundamental insight of the philosopher who turned empathy (a term developed in German nineteenth century aesthetics) into a psychological concept. Empathy, Theodor Lipps (1903) claims, is a form of perception; but contrary to what Max Scheler (1913/1979) later claimed, it isn't, according to Lipps, conatively neutral. Scheler argued that the fact that one person empathizes with another does not, in itself, say anything

about his practical attitudes; a sadist may empathize with a person whose suffering he enjoys and be motivated to increase that suffering, while a sympathetic person will rather be moved to alleviate the pain. Lipps, by contrast, argues that there is something like a sympathetic impulse involved in empathy, and that this is basic for our understanding of other people's minds (a claim that fits seamlessly with Michael Tomasello's (1998) view that toddlers grasp other people's intentions long before they have a theory of mind). Empathy is, according to Lipps, 'internal co-action', a claim which is very much in line with current simulation theory and the role of mirror neurons. This is, of course, not to deny that Schelerian 'antipathetic empathy' is possible; but the fact of the matter is that the two cases of sympathetic and antipathetic empathy are not on a par; while there needs to be some antipathy at work in the unsympathetic cases (such as the desire to see the other person suffering), no additional pro-attitude beyond the mere fact of empathy is necessary to explain the sympathetic effect. Consider the case of an elderly person struggling to lift her suitcase onto the luggage rack. There is an immediate *impulse* to lend her a hand, and current neurological research seems to suggest that it is in the light of this impulse that our understanding of what she is trying to do comes about.

This is not to deny that this impulse cannot be suppressed and that agents can acquire a disposition to remain passive in such situations. In fact, suppressing one's sympathetic impulses is an important part of the process of socialization. This is due to the fact that, while the first interactions in which a child engages are cooperative in nature (mother-child interaction), competitive interactions become prevalent in later stages of a person's life. While the empathic impulse provides the motivational steam for success in cooperation, one must be able to suppress that impulse – i.e. to take one's mirror neurons entirely offline, as it were – in competition. Successful competitive behaviour requires that an agent be aware of his competitor's motivations, not so as to cooperate but rather so as to use this information to further his own anti-pathic agenda.

Moreover, and more interestingly, even some civilized cooperative forms of interaction require the agent to suppress her empathic impulses, a point of special importance in that it helps to dispel the view that everyday altruism might be purely norm-driven. It is true that in most cases (such as the cases of everyday altruism quoted above) action on emphatic impulses is supported by the rules of politeness and proper conduct, which may lead some into thinking that the phenomenon in question is really a matter of manners rather than motivation. Interestingly, however, there are many cases where the emphatic impulse is *in conflict* with the norms of proper conduct. This is especially true in areas where a person's autonomy is at stake, and especially also respect for her agency, whether because of the person's handicaps or because she

is a child and needs to be given the opportunity to exercise her own agency without being interfered with. Generally speaking, this is true wherever it is not only important that people's goals are achieved, but also that they achieve their goals *themselves*, without outside interference. A person who cannot suppress her empathic impulse would be a rather bad parent, giving her child no room to develop a sense of his own agency. And as a perhaps even more obvious example, politeness strictly requires of us to suppress the impulse to finish a sentence in which a person struggling with stuttering is stuck. In many cases, the empathic impulse is supported by social norms of propriety; in other cases, it clearly is not.

Thus the ability to suppress one's empathic impulses is an important part of the process of socialization, both with respect to competitive success and conformity with the social norms of propriety. To the degree that such a disposition is acquired, empathy becomes conatively neutral, and such agents may need an extra pro-attitude to become active (such as the desire to be polite, or some other self- or other-directed motive). But this structure of conatively neutral empathy should not be mistaken for the basic mode.

The empathic impulse is the most fundamental form of philosophical altruism. It is not, however, commonplace to be pushed to act by motivational impulses, be it one's own urges or what one perceives to be some other person's goal (remember the Kantians' worries). In standard cases of action, the agent is not entirely passive with regard to his or her motivational base. Standard action is *deliberative*; the role of deliberation is to identify one's reasons for action by making them effective (e.g. Searle 2001). One might be tempted to see deliberation as a process by which empathic impulses are ruled out as proper reasons for action, and by which motivational autarky is achieved. After all, how could the fact that another person wants to A be a reason for a deliberatively rational, non-impulsive person, without that other person's desire having some value in the light of the agent's own pro-attitudes? The standard view seems to be that for such agents empathy has to be conatively inert; empathy *informs* such agents of other people's motivations, but does not, in itself, provide them with a reason to act – or so it seems. Without launching into a conceptual analysis of empathy here, it seems, however, that empathy plays the exact same structural role in practical deliberation with regard to other people's pro-attitudes as the agent's self-awareness does with regard to his own desires. Given this analogy of the agent's awareness of her own desires and her empathetic awareness of other people's it is not at all obvious, why the agent's own desires should play a structurally different role in her practical deliberation than those of another person. In other words, that a person should consider another person's desire as a reason for action is no more mysterious than that she should treat any of her own desires in that way.

Another term that is sometimes used in the literature to describe this structure is *interpersonal identification* (which Sigmund Freud (1921/2005) classifies as the most fundamental mode of affective attachment between persons). There is an air of paradox about this term since to identify x with y is to judge that x is y , which is at odds with the claim that identification may be interpersonal rather than intrapersonal. We seem to be coming dangerously close to Schopenhauer's claim about the ultimate unity of all persons here; but we need not go that far. Suffice to say that person A identifies with person B to the degree that B's pro-attitudes are included in A's class of possible reasons for action. Contrary to what Schopenhauer seems to think, the fact that the classes of reasons for action may overlap (or even be one and the same, where identification is total) does mean that, on some deeper metaphysical level, a distinction between persons does not exist. The identification is between *different* people, and yet they do not take their own motivational states to be the only ultimate reasons for action, but rather extend the class of potential reasons for action beyond their own psychology.

In real cases, identification is *selective* – a person identifies herself with some people, but not with others – and it is *a matter of degree*: a person may include some of the other's desires in the class of her possible reasons for action while excluding others, and she may do so to a greater or lesser degree. Thus the question is: how is the range of people with whom an agent identifies, and the degree to which she does so, determined? It seems to me that the term 'empathy' as well as Freud's claim that the kind of interpersonal relation that is established in identification is *affective* points toward the right answer. Empathy and identification are *affective* attitudes, and it would be interesting to examine other-directed emotions in terms of how exactly they lead those having them to include the motivational and conative states of the others to whom they are directed in the base of their own practical deliberation. It is likely that such attitudes as trust and respect fulfil this role differently from friendship or love. In this view, such emotions should not be seen as motivational states in themselves; rather, they should be seen as modes of identification, i.e. ways in which an agent's class of possible motivating reasons for action is extended beyond her own subjective motivational set. The question of whose motivations provide an agent with reasons for action, and to what degree they do so, is basically a matter of the affective attitude an agent has towards other persons. In concluding this part of the discussion, it might be worth mentioning that this reading fits rather nicely with Auguste Comte's original idea concerning the nature of altruism. According to Comte, altruism should be neither viewed as a way of thinking, nor as a way of acting. Rather, Comte situates altruism in the third of the domains he distinguishes, i.e. the sphere of *sentiments*, the affective sphere (Comte 1851: 694ff).

6. PHILOSOPHICAL EGOISM AS A VALUE

How is the case where an agent forms an intention on the basis of another person's desire, simply because she is identified with that person, different from the case where she makes the other person's goals her own? Having made another person's goal, or desire, one's own means having a motivating other-directed desire to fulfil the other person's wish. This means being able to account for the degree to which other people's desires influence one's course of action in terms of one's own motivational agenda. Using the terms introduced above, a person who does not let herself be influenced by what other people want, apart from making other people's desires her own, is *motivationally autarkical*. If she complies with another person's demands, or lends another person a hand, or gives in to some empathetic impulse, she does so only if and to the extent that this is what she wants, in the motivational sense of the term. She draws entirely from her own motivational resources. Acting on other-directed motivating desires might *presuppose* some degree of identification; however, it is more than that. There is a sense in which such a person *volitionally endorses* the other person's attitude that goes beyond mere identification. Identification does not just happen to her; rather, she *wants* it, she is *motivated* to be so identified. When she is moved to act, she does so not only because of another person's desire, but because this is what her own desires demand.

Such a person is fully self-reliant, and motivationally autarkical. Even while acting with devotion in the interest of others, with no goal other than to promote their well-being, the Nietzschean 'I will' is written in capital letters over her actions. There is a sense in which motivational autarky captures our sense of what it means to be a fully developed person. Such a person should not do anything for the simple ultimate reason that this is what another person (with whom she identifies herself) wants; rather, she should do so only insofar as this is what she *herself* wants. To be a fully developed person requires a sort of responsibility, i.e. an ability to account for one's actions in terms of one's own motivations. Only such a person has the 'motive principle', of which Aristotle speaks in his reflections on action, fully within herself. Never would she have to resort to external factors in basic motivational explanations of her actions. In the last resort, she is bound by her own will only.⁶

⁶ It would be interesting to see if philosophical egoism might be at the heart of what seems attractive in Max Stirner's normative ideal, presented in *The Ego and His Own* (1845/1995), especially since the label 'philosophical egoism' is often associated with his views. Stirner's work, as well as his critics', is notoriously vague with regard to the distinction between psychological and philosophical egoism, making it difficult to ascertain what Stirner's position really amounts to. In some passages, he seems to reject philosophical egoism as a *structural* feature of action. This is especially obvious where he speaks of people's being 'possessed' by motives of which they are not the owners, or a will which is not their own,

Philosophical egoism is certainly a cultural ideal, and it is closely intertwined with some of the thickest notions of our moral vocabulary, such as personhood, autonomy, and responsibility. But people very often act on other people's desires without having set up a motivational agenda of their own, simply because they find themselves trusting, loving, respecting other people, perhaps even against their will, or because they find their reasons for action permeable to other people's desires in some other way. Such people may fall short of our full-fledged notion of personal identity, but even if we disapprove of such behaviour (there seem to be opposing views),⁷ there is no reason to ignore its existence. In received theory, there is a tendency to mistake philosophical egoism for a structural feature of agency rather than taking it for what it really is: a cultural ideal of personal development. This is particularly obvious in intentionalistic readings of rational choice theory, where individuals are taken to be motivationally autarkical beings. I have argued in this paper that this is mistaken. Most people are not full-fledged philosophical egoists, and hardly anyone has always been one.

REFERENCES

- Andreoni, J. 1990. Impure altruism and donations to public goods: A theory of warm glow giving. *The Economic Journal* 100/401: 464–477.
- Batson, C. D. 1991. *The Altruism Question: Toward a Social-Psychological Answer*. Hillsdale, NJ: Lawrence Erlbaum.
- Becker, G. S. 1986. The economic approach to human behavior. Reprinted in *Rational Choice*, ed. J. Elster, Ch. 4. Oxford: Oxford University Press.
- Comte, A. 1851. *Système de Politique Positive ou Traité de Sociologie*. Vol. 1. Paris: Librairie de L. Mathias.
- Davidson, D. 1963. Actions, reasons, and causes. *The Journal of Philosophy* LX (23): 685–700.
- Fehr, E. and U. Fischbacher 2003. The nature of human altruism. *Nature* 425: 785–791.
- Feinberg, J. 1958/1995. Psychological egoism. In *Ethical Theory: Classical and Contemporary Readings*, 2nd edn, 62–73. Belmont: Wadsworth.
- Frankfurt, H. 1971. Freedom of the will and the concept of a person. *Journal of Philosophy* 68: 12–35.
- Freud, S. 1921/2005. *Massenpsychologie und Ich-Analyse*. Frankfurt: Fischer.
- Garfinkel, A. 1981. *Forms of Explanation. Rethinking the Questions in Social Theory*. New Haven: Yale University Press.
- Hollis, M. 1998. *Trust within Reason*. Cambridge: Cambridge University Press.

making one think that the egoism of 'full self-possession' that he ends up recommending in the third part of his work might really just be philosophical rather than psychological. However, this expectation is frustrated in most passages, as Stirner clearly argues for *psychological* egoism as a normative ideal.

⁷ For an alternative normative account of selfhood, cf. the work of the French phenomenologist Emmanuel Lévinas. According to Lévinas, being 'possessed' by the needs of others in a way that expropriates the agent of his self-ownership is no deficient mode of selfhood, but rather at the foundation of ethical character (cf. Lévinas 1991).

- Kitcher, P. 1998. Psychological altruism, evolutionary origins, and moral rules. *Philosophical Studies* 89: 283–316.
- Lipps, T. 1903. *Grundlegung der Ästhetik*. Hamburg/Leipzig: Voß.
- Lévinas, E. 1991. *Entre nous. Essais sur le penser-à-l'autre*. Paris: Bernard Grasset.
- Nagel, T. 1970. *The Possibility of Altruism*. Oxford: Clarendon Press.
- Paprzycka, K. 2002. The false consciousness of intentional psychology. *Philosophical Psychology* 15/3: 271–295.
- Pettit, P. and M. Smith 2004. Backgrounding desires. In *Mind, Morality, and Explanation*, ed. F. Jackson, P. Pettit and M. Smith, 269–293. Oxford: Oxford University Press.
- Scheler, M. 1913/1979. *The Nature of Sympathy*. Translated by P. Heath. Hamden, CN: Shoe String Press.
- Schopenhauer, A. 1840/1995. *On the Basis of Morality*. Translated by E. F. J. Payne. Providence, RI: Berghahn Books.
- Schopenhauer, A. 1849. *Die Welt als Wille und Vorstellung*. Wiesbaden: Brockhaus.
- Schueler, G. F. 1995. *Desire: Its Role in Practical Reason and the Explanation of Action*. Cambridge, MA: MIT Press.
- Searle, J. R. 2001. *Rationality in Action*. Cambridge, MA: MIT Press.
- Sen, A. K. 1977. Rational fools. A critique of the behavioral foundations of economic theory. *Philosophy and Public Affairs* 6: 317–344.
- Sen, A. K. 1985/2002. Goals, commitment, and identity. Reprinted in *Rationality and Freedom*, 206–225. Cambridge, MA: Harvard University Press.
- Simmel, G. 1892. *Einleitung in die Morawissenschaft*. Vol. 1. Berlin: Hertz Verlag.
- Sober, E. and D. S. Wilson 1998. *Unto Others*. Cambridge, MA: Harvard University Press.
- Stirner, M. 1845/1995. *The Ego and His Own*, ed. D. Leopold. Cambridge: Cambridge University Press.
- Tomasello, M. 1998. *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.
- Trivers, R. 1971. The evolution of reciprocal altruism. *Quarterly Review of Biology* 46: 189–226.
- Williams, B. 1981. Internal and external reasons. Reprinted in *Moral Luck*, 101–113. Cambridge: Cambridge University Press.