Altruism and the Experimental Data on Helping Behavior

Stephanie Beardman

Abstract Philosophical accounts of altruism that purport to explain helping behavior are vulnerable to empirical falsification. John Campbell argues that the Good Samaritan study adds to a growing body of evidence that helping behavior is not best explained by appeal to altruism, thus jeopardizing those accounts. I propose that philosophical accounts of altruism can be empirically challenged only if it is shown that altruistic motivations are undermined by normative conflict in the agent, and that the relevant studies do not provide this sort of evidence. Non-normative, purely causal, psychological factors would be empirically relevant only if the notion of altruism is broadened to include the requirement that one recognize certain situations as calling for altruism. But even in that case, the relevant studies are not designed in such a way that could threaten philosophical theories of altruism.

Keywords Altruism · Helping behavior · Darley · Campbell · Good Samaritan study

Moral attributions play a large role in much ordinary explanation of human behavior. We say someone was motivated by honor, duty, kindness, goodness, compassion, or that she acts with integrity or because she is virtuous. We also attribute immoral motivations, such as bad or evil intentions, spitefulness and malice, self-despising feelings, or desires to be cruel, unkind, or to humiliate. Mixed or non-moral motivations can be explanatory as well, as happens when we say someone acts from unreasonable, unfair, selfish, or ideological inclinations. We can understand the act better when we understand the motive of the actor. This paper will focus on how psychological studies on the circumstances affecting the likelihood of certain behavior might threaten such normative explanations. In particular, I am interested in normative explanations of helping behavior—that is, in explanations that attribute altruism as a moral motivation.

S. Beardman (✉)
Department of Philosophy, Barnard College, Columbia University, 3009 Broadway, Milbank 326, New York, NY 10027, USA
e-mail: sbeardman@gmail.com

There is a large body of evidence suggesting that situational factors have a great, and systematically underestimated, influence on human behavior. Work spanning 50 years in a branch of social psychology called attribution theory shows that the environmental, circumstantial features of a situation are far better predictors of certain behaviors than are the attributes of the persons performing those behaviors. The probability of getting help from a bystander goes up as the number of other bystanders decreases. The fewer the siblings, the more likely one is to receive a kidney from a sibling. Noise from a lawnmower dramatically decreases the chance that a person with a broken arm will be helped when a bunch of books slide out of the box he is carrying. These examples represent a tiny sample of the studies on prosocial, or helping, behavior. The literature is rich with many interesting results about the usefulness of situational features in predicting and explaining other types of behavior as well.[1]

We already believe, independent of scientific experiment, that certain circumstances are relevant to the likelihood of helping behavior. For example, if the costs of helping are very high, then we can easily predict and explain—and, in some cases, morally excuse—the lack of helping. The striking thing about the work in attribution theory, however, is that it highlights situational features that we take to be morally irrelevant to the performance of the act in question. Plenty of studies suggest that, often enough, costs are not part of the right explanation; people are inhibited from doing the morally right thing by more than high costs. Finding a dime in a phone booth, for example, would not seem to change the costs and benefits of helping someone who has just dropped a folder of papers nearby; yet it does change the likelihood that helping behavior will occur. It seems that morally irrelevant circumstantial features diminish the likelihood of being altruistically motivated.

What are we to make of these studies? Philosophical accounts of altruism attempt to identify what it is about morally worthy helping behavior that makes it morally worthy. By suggesting competing, non-normative explanations, the situationist studies on helping behavior have been interpreted as bearing on these normative accounts of altruism. Could the great number of morally irrelevant circumstantial features that causally influence the possibility of altruism shed any light on the nature of altruistic motives? Does it undermine the existence of the sort of altruistic motive that is posited by moral philosophers? In the next section, I reconstruct the argument that such studies pose a threat to philosophical accounts of altruistic motives. After addressing the larger issue of whether the heavy influence of situational factors poses a threat to normative explanations of behavior in general, I argue that the empirical work that underlies situationism does not have much to do with the normative issues at the heart of philosophical accounts of altruism, because the studies themselves are not designed to be able to do so.

1 A Sample Study and the Argument Against Altruism

John Campbell has presented a dilemma for theorists who offer accounts of altruism.[2] Either they do not present empirical hypotheses about actual helping behavior or they do. In the former case, the right account of morally good helping behavior is held to be completely independent of what people are actually like. The correct account of morality would have nothing to do with

---

[1] See Ross and Nisbett (1991), Mischel (1968), Doris (2002) for the details of these and other relevant studies.
[2] "Can Philosophical Accounts of Altruism Accommodate Experimental Data on Helping Behaviour?" (Campbell 1999). Unless otherwise noted, all page references are to this text.

explaining why people help others. On this view, experimental data about what explains helping behavior could not possibly illuminate what is good about morally worthy helping behavior. Moreover, it may in fact turn out to be the case that no instance can be found of morally worthy helping behavior—however many cases of helping behavior occur. Campbell stays neutral on the merits of this sort of view, and I do not want to pursue it here. Most theorists are not so sanguine about the possibility of there being no morally good acts. In particular, they hold that there are morally good acts of helping behavior. I, too, am interested in philosophical accounts of altruism that are at least partly meant as empirical theses, in that they purport (in part) to provide the correct *explanation* of a not-insignificant portion of actual helping behavior. On these views the motive of altruism is capable of causally explaining behavior, and it is this feature of the explanation that makes the behavior morally worthy.

Campbell argues that those who take this horn of the dilemma fall prey to the 'fundamental attribution error', which involves grossly underestimating the effect of situational determinants on behavior while overestimating that of personal or dispositional features.[3] In appealing to altruistic motivations to explain action, such theorists overlook the large causal role played by the situation and people's construals of it. Their confidence in the existence of altruism is misplaced; in fact, "given various empirical data, it is questionable whether such [altruistic] motives do play a role in the explanation of helping behavior" (33). As a result, Campbell suggests that the data "require rejection of various philosophical accounts of altruism" (26). The following are among the accounts of altruism he mentions: Kantian analyses according to which altruism is a matter of acting from duty, or of having an integrated and authentic self, or of having a conception of oneself as just one person among other equally real selves (cf. for example Nagel, 1970); Humean accounts that take altruism to be a matter of feeling sympathy or other compassionate emotions (cf. for example Slote, 2001); Virtue-ethical accounts according to which altruism is tied to having a virtuous perception of the situation that stems from having a good character (cf. for example McDowell, 1979).

These accounts posit altruism, a type of motivation, as a state that can be attributed to an agent to help explain certain acts. Note that 'altruism' here does not refer to the notion—frequently discussed by evolutionary psychologists—of helping another at the expense of one's own interests. In addition, it does not require that one succeed in helping another. Depending on the philosophical theory in question, it may require that it issue forth from a virtuous, or at least altruistic, character. *But it need not*; altruistic motives are not always tied to broad dispositions. At a minimum, altruistic motivation involves the desire to improve another's situation, though it may do so only derivatively.[4] It will be understood differently by different moral theories. Regardless of their differences, however, these

---

[3] Impressed with the lessons of the fundamental attribution error, some philosophers argue that the error has significant consequences for ethics. They have argued mostly that the data undermine the notion that we have robust, stable character traits of the sort needed for certain forms of virtue ethics. Cf. Doris (1998, 2002) and Harman (1999). Attribution theorists themselves have also argued that their work has moral implications of various sorts, especially concerning the attribution of moral responsibility. See Schoeman (1987) for some discussion of their work. Finally, philosophical accounts of altruism in general have also been supposed to be undermined, and this is what I will speak to.

[4] Accounts of morally worthy helping behavior will generally, I think, take the desire to help another to be an ultimate, and not merely instrumental, desire. Altruism in this sense contrasts with psychological egoism, the theory about motivation that holds that the only thing that is intrinsically desired is something such as one's own perceived well-being. I leave to the side here the thorny issue of whether and how the distinction between instrumental and intrinsic desires is amenable to empirical investigation. The distinction is not relevant to the sort of argument considered below. (I say a few more words about the egoism/altruism debate when discussing what is at stake in Section 2.)

theories invoke altruism not merely to help explain, but to identify what makes helping behavior morally worthy. Thus the notion is, importantly, a normative one. As such, it is not merely a cause: a person who acts from the motive of altruism will be acting for altruistic reasons, which count as normative reasons. In an effort to remain neutral among the various possible accounts, one might say that altruism is a response to a normative reason to help others. I do not mean here to presuppose any particular theory of motivation.

Campbell argues that the evidence suggests that—as an empirical hypothesis of what causes helping behavior—each of these philosophical accounts must be false. Rather than survey a large group of studies, he focuses closely on a psychological study that has gained some prominence in the literature. This is the 'Good Samaritan' study of Darley and Batson (1973). It goes as follows. The subjects were students at the Princeton Theological Seminary who each volunteered to give a short sermon (some, on the parable of the Good Samaritan). On the way to their talk, they passed a man slumped in a doorway who coughed and groaned. Some of the subjects did help, most did not. Earlier, each had filled out a 'religious personality' questionnaire designed to measure the motivation for their interest in religion. They were then instructed to go to a room in a building close by to give their sermon. Some subjects were told that they would have to hurry because they were late, others were told that they were right on time and could go right over, and a third group were told that they were early but wouldn't have to wait too long once they got there. With which of these variables—subject matter of the sermon, religious orientation, degree of hurriedness—was the helping behavior correlated? The data show that helping behavior was highly correlated only with degree of hurriedness. In fact, 63% of those who were least hurried stopped to help, 45% of those in the intermediate group, and only 10% of those who were most hurried. So 90% in that last group did not even alert another person to the plight of the man they had just passed by. These are, in the briefest terms, the relevant results of this well-known study. The question before us is whether this study helps support a challenge to the existence of genuine altruism.

What should we make of the Good Samaritan study? Here is one answer, proposed by Campbell.

> Upon encountering a distressed person, whether an individual will assist basically depends on the individual's degree of hurriedness to do something else. It does not depend on the individual's character, e.g., on whether the individual is compassionate or callous. Nor does it depend on the presence or absence of a sense of duty to assist those in need, or on an integrated authentic self, or on perceived self-interest in helping, or on a conception of oneself as one person amongst others equally real, etc. We should just junk views that purport to account for helping behaviour in those sorts of terms. We should junk them because they are empirically false accounts of the basis of helping behaviour. (31)

The claim is that the large number of studies like that of the Good Samaritan shows that normative accounts of helping behavior should be rejected. As I understand it, the argument for adopting this proposal is supposed to go something like this. Suppose people were capable of having altruistic motives. It would be surprising if those motives were so fragile that they explain behavior only in very rare and special circumstances, such as when one is unhurried, there is no loud background noise, there are no bystanders, and so on. And yet, the data seem to show just this. They show that altruistic motivations, if they exist, can be undermined by seemingly trivial situational forces that overwhelm the motivation (such as whether people are hurried, or whether other people are present, or whether one has previously been exposed to similar cases, and so on). It is a condition on behaving

altruistically that subjects perceive an opportunity to benefit another. And it would not be surprising to learn that time pressures may affect what people notice.[5] (I address the fact that situational features also influence one's cognitive/perceptual capacities—they diminish the likelihood that certain circumstances will be perceived—in Section 4.) But enough non-helpers in this study *did* perceive the distressed person (some appeared agitated upon entering the building; some recalled their inaction later when asked to identify an occasion when someone needed help).[6] Being in a hurry affected their capacity to act. Should we count a motivation that is so *easily* overridden genuine altruism? Surely, the idea goes, that is too feeble a motive on which to hang a whole moral theory. This reasoning captures much of the implicit thought behind the large popular reaction to these surprising empirical results, and I think it is essentially what Campbell proposes. Here's how he puts the crux of the challenge:

> The difference in the probability of helping is explained by the fact that the first set is in a bit of a hurry to do something else. Isn't that a pretty trivial thing? Remember, they weren't off to perform emergency brain surgery. They were off to record some relatively unimportant talk and it was only a matter of being a few minutes late to do it. The subjects had volunteered to participate in 'a study on religious education and vocations' and they had been asked to record a brief talk to supplement a questionnaire that they had completed. Running a bit late to record such a talk seems inconsequential compared to taking a moment to check on someone in severe distress.... (41)

The issue seems to be that considerations that we—and presumably, the subjects, too—take to be morally irrelevant so easily, and apparently unconsciously, prevent or override the morally good motivation to help another.

2 What's at Stake: Normative Explanations

Before responding to the challenge, it's worth emphasizing that the stakes in this debate are much higher than has been generally made explicit in the literature. The vast majority of the literature, especially over the last 10 years—too large to list here—has focused on the implications of situationism for virtue ethics. But in fact, if it's true that the sort of empirical study presented here poses a threat to philosophical accounts of altruism, then it poses a threat to all normative explanations. By 'normative explanation' I mean an explanation of behavior that makes moral attributions, of the sort illustrated at the outset of this paper. Such explanations invoke normative terms, such as having mental attitudes that are morally or non-morally good, rational, and the like. If Campbell is right, then not only might altruism not explain action, but neither would morally good (or bad) motives of any sort explain anything.[7] To see why, one has only to see what all philosophical accounts of altruism have in common.

---

[5] Note that the issue is not just that of attention, but also concerns disentangling the effects of the ambiguity of the situation (the distressed confederate may have seemed menacing, for example). Perhaps, also, there is a 'diffusion of responsibility' effect at work here, if the subjects could reasonably expect that other potential helpers would also pass by the same area.

[6] Moreover, there exist plenty of other studies with similarly surprising behavioral results, in which it is not plausible to think that perception is impaired.

[7] Of course, there are philosophers who have suggested, for various reasons, that we should do away with such explanations anyway.

As noted above, philosophical accounts of altruism do not all share the view that there are character traits. Acting from the motive of duty, for example, does not obviously require that one have a stable disposition to do so. Nor does being moved by sympathy require such traits. Rather, when committed to explaining behavior, philosophical accounts of altruism share the view that there are certain types of motivations—connected in the right ways with duty, authenticity, virtue, sentiments, or what have you—that make helping behavior morally worthy. If this feature, the positing of altruistic motives, is found to be problematic or illusory, then it seems the problem could not be altruism per se.

Rather, the problem would have to stem from the fact that these motivational states are *normative*, that they are morally good (or bad) motives. If the evidence calls into question the existence of altruistic motives, it would seem to call into question the normativity of all motives. So, in suggesting that there are no altruistic motives, the studies in question would thereby challenge also the existence of cruel, selfish, and other such normatively laden motives. Campbell comes close to recognizing this point when he writes,

> the basic problem here is not restricted to the explanatory power of altruistic forces but is a more general one. It is not just the idea that there are powerful altruistic forces that may be difficult to reconcile with this experimental data. The idea that there are powerful forces toward indifference [or callousness] may be equally difficult to reconcile with the data. If the idea that people are compassionate is difficult to reconcile with the fact that they pass by when they are in a bit of a hurry, the idea that people are callous and indifferent is equally difficult to reconcile with the fact that they help when unhurried. (42)

Campbell's explanation is that the problem lies in attributing behavior to qualities that "cannot be so easily undone or undermined" (42), such as genuine compassion, character traits, and the like. For example,

> …the problem is in thinking that it is persons' character (whether compassionate or callous) that determines whether they engage in helping behaviour, rather than their degree of hurriedness. A similar point applies to the other philosophical theses outlined earlier. For example, there is a problem in thinking that it is the nature of the self (whether integrated or conflicted) … [or] that it is a sense of duty to help (whether present or absent) that determines whether persons engage in helping behaviour, rather than their degree of hurriedness. (42)

What is particularly arresting about Campbell's argument (though it seems he may not appreciate just how far-ranging the implications of this are) is that it raises the possibility that there is an empirical threat to all theories of morally good helping behavior, and not just to virtue ethics.[8] These studies do not present a particular threat to the existence of robust and stable character traits, but pose a deeper challenge: they call into question core assumptions about morally charged motives (good, bad, or indifferent). In doing so, they threaten normative explanations in general. Identifying *motivations* as the target of the empirical challenge implicates the viability of all normative explanation of behavior.

Note that this is not the familiar conceptual point about the difficulty of distinguishing altruistic from, say, selfish motives. Much has been written on the question whether people's altruistic desires involve taking another's well-being as an ultimate end, as

---

[8] Sabini and Silver implicitly recognize the same point in passing, when they claim that the standard situationist interpretation of certain social psychological studies "does impugn all forms of ethics," and that it "denies the idea that behavior follows from beliefs, desires, and values" (2005, 536, fn. 5).

opposed to an instrumental means to some ultimate egoistic end. Whatever one thinks of the empirical tractability of that question,[9] it should be clear that the egoism/altruism debate is a more difficult one than the present question whether people are often enough motivated to help others (whether ultimately egoistically or not) in a way that is not situationally determined, or than the question whether such motivation enters often enough into the best explanations of people's behavior. Even if it turned out that nearly all the seminarians stopped to help, and that in general helping behavior is widespread in a way that is not situationally determined, that in itself would not help one way or the other to settle the traditional debate between egoists and altruists. Yet, though they would not bear directly on that question, such facts would, of course, help to block Campbell's present worry that there is too little helping behavior that isn't situationally determined to make what might be motivating that behavior a centerpiece in ethical theory.

Furthermore, in arguing about what is at stake I am not raising a *general* conceptual worry concerning the difficulties of identifying which features of a situation are relevant to eliciting particular motives. I do not deny that environmental conditions such as those created in the Good Samaritan study are obviously relevant to studying helping behavior, and by extension, possibly the motives behind such behavior—though much will depend on the methodological adequacy of the experimental design, as I explain below.

3 Responding to the Challenge

Keeping in mind that the study presented here is representative of a very large group of similar studies, the question is what to make of the fact that most of the time, people in general seem motivated by morally trivial concerns; seemingly insignificant situational factors have a disproportionately great effect on behavior.

The first thing to say in response is that doing the right thing can be difficult. None of the philosophical accounts are committed to the claim that altruism is either easy or very common. It may require being attentive and, in one way or another, inured to the influence of a plentiful array of distracting factors. So, one could say, the studies here show that performing an altruistic act while challenged by situational features is a rare and difficult thing, but not impossible, as the 10% of hurried subjects show. Campbell does not seem to address this (fairly common) response head on, but what he says at various points suggests that he is aware of and open to it. He acknowledges that those who helped despite being hurried may have "significant altruistic impulses" (43). He insists that whether they do is an empirical matter; it would need to be supported by evidence that those who do help are in some empirically identifiable way—he suggests personality assessments—significantly different from those who do not help. I would add to this that it's important, of course, that these psychological differences be of the right sort. For example, perhaps it is difficult to make some people respond to being under time pressure; we should keep in mind that the experimenters may not have succeeded in making this 10% group of subjects feel hurried.

---

[9] John Doris thinks that it would be, as he puts it, a "serious mistake" to interpret the empirical evidence as counting against the existence of altruism. However, his reason for thinking this is that he doubts that "questions about the possibility of altruism admit of empirical resolution, since the issue concerns what sort of motivations should be counted as altruistic, and this is substantially a conceptual difficulty" (2002, 35). Dale Jamieson (2002) argues that, on the view that desires are internal representations, one can never be sure that a particular desire takes others' welfare to be ultimate. However, he thinks that there is empirical evidence that renders the view that people have psychologically ultimate altruistic desires more plausible than the alternative.

Thus one can interpret the studies as showing only that there are many pitfalls that can prevent altruism.[10] This seems the most direct route to arguing against Campbell's claim that philosophical accounts of altruism are empirically false. Moreover, it seems right to recognize that morally good acts can be difficult in surprising ways, and that they require being impervious to an incredible variety of morally irrelevant considerations, most of which one is not aware.

However, even if one endorses this response, an empirical question remains. The response just presented makes one wonder if *anything* would count as an empirical challenge to the philosophical accounts. Indeed, I suspect that many who would offer this response to Campbell—that being altruistic may be exceedingly difficult in certain circumstances—might be tempted, in the end, to relinquish the empirical commitments of the moral theories in question. If the normative concept of altruism is not required to explain any action, they might want to say, so be it. Being good is hard. Such a position embraces, perhaps happily, the first horn of the dilemma described above, that philosophical accounts of altruism are not meant to offer empirical hypotheses. I do not want to take that route here.

Instead, I'd like to explore a different response to the purported empirical challenge to philosophical accounts of altruism. This response offers an interpretation of the data that is compatible with the one just described (that it may be uncommon and difficult to be altruistic). However, it does not require it, and it has the important benefit that it proposes an empirical criterion for the adequacy of the philosophical accounts. In short, the proposal is this. A challenge to these accounts can be posed only if it can be shown that (morally trivial) *reasons* outweigh people's motivation to help others. The evidence, such as it is, so far yields no reason to think that the subjects in the study did have such reasons. In what follows, I'll develop this thought.

Whether a philosophical account of altruism grounds that motive in a person's desire to help others, or in a person's recognition of the fact that doing so is good, or in having a virtuous character, altruism is nothing if not a response to a reason to help others.[11] The motivations appealed to in philosophical accounts of altruism must have something to do with an agent's own reasons for acting—otherwise, the motivations, and altruism itself, would not be normative. (I am not claiming that the agent must be explicitly aware of these reasons.) This is why the empirical data on helping behavior do not clear the way for a straightforward attack on or challenge to the possibility of altruism. They do not establish that the subjects were what one might describe as 'normatively moved' by morally trivial concerns.

The existence of altruism cannot be undermined by evidence that non-reason-giving considerations diminish the frequency of helping behavior. Some support for this can be found in our common practices of attributing motives. Consider the fact that for the vast majority of people during the vast majority of the time, one usually is not motivated by altruism. This is often because one is engaged in activities not related to helping others. Do we want to say that the lack of an altruistic motive in such cases has a bearing on our theories of altruism? Of course not. It takes more than the mere fact that one is moved by factors that are not relevant to altruism to establish that there is little to no altruism.

The basic point might be brought out by considering the total irrelevance of exculpatory reasons in the context of the Good Samaritan study. For example, one might be excused for not helping another if one's life or limb is under threat for doing so. In contrast, the

---

[10] There is the further, distinct, question of how to evaluate the actions of those subjects who do not help, and of whether they are blameworthy for failing to help. I will return to this below; my own interpretation of the study does not entail that they are. Some may be epistemically blameworthy (because they didn't notice) and others morally blameworthy (because they noticed but did not care), and yet others blameless in both senses.
[11] There may be a problem with this formulation, connected with my use of 'reason'-talk. As such talk is ubiquitous, most should not have a problem with my reasoning on that account.

comparatively trivial consideration that one will be made late for an unimportant presentation is not exculpatory. But *this* observation is not automatically relevant to understanding the subjects here. The usual assessments regarding whether people are excused because of the risk of high costs *do not apply* in this case. This is because the consideration that one will be made late must be *their* reason—and not only that, it must be their reason *for not helping*. Otherwise, in this case, it has merely explanatory force vis à vis the not helping. It has no normative status in explaining why there was no helping. Candidates for exculpatory reasons are reasons that the agent can (and in the circumstances, does) weigh against the reasons she may have to help another. But the consideration of being in a hurry is in this case not a reason in the relevant sense at all. After all, there is a distinction between being in a hurry and having a (subjective) reason not to be late. A subjective normative—altruistic—reason must be on the table before we can even speak of whether such trivial reasons morally excuse the behavior in question.[12]

This study has not captured anything like an operative reason of this sort. This is not a defect of the study; it was not designed to do so. The study randomly assigns individuals to different experimental conditions and compares the average response of those groups. The observed differences in the helping response cannot be accounted for by differences in the average motivation, personality, or other traits in the groups, since the groups were randomly selected. Thus, the design of the experiment itself blocks appeal to individual differences as an explanation of the behavior observed in the experimental context. The differences among the hurried/less hurried/not hurried groups cannot be explained by individual differences, and in fact there is no way to discern these, given the study's design. All that the experimental method can tell us here is the effect of the variable manipulated (hurriedness) on average rates of helping among between-cell groups (the dependent variable).

This means that it would be incorrect to attribute to all non-helping subjects a subjective normative reason to help another that they ignored or failed to be moved by, because their concern to arrive at a scheduled appointment on time (normatively) outweighed their reason to help someone, or to say that being hurried provided a (subjective) reason to ignore someone in need. Being hurried *was*, most likely, a reason for hurrying to give the talk; but it was not a reason to refuse to help. It is exactly for this reason that it is in fact more natural to say that they failed, rather than that they refused, to help—though even this is pure speculation, for the reasons just explained.

There are two ways that altruism may be thwarted: one is as a result of some type of normative conflict and the other is because it was diminished or disabled by purely causal psychological factors. The latter is a case of what I call 'non-normative overriding'.[13] We

---

[12] In the Biblical parable of the Good Samaritan, of course, the priest and the Levite did have such reasons clearly present to them. In the story, it's clear that each man came directly upon the hurt and wounded man lying by the road, saw and understood the bad state he was in, and decided to cross the road and pass by on the other side. They are each motivated anti-altruistically—and that they are so motivated is essential to the power of this parable. It highlights the proper moral understanding of what the Samaritan did in contrast to the actions of the priest and Levite (Luke 10: 27–37, King James version).

[13] Schoeman (1987) seems to invoke this same distinction when discussing what makes certain decisions hard. "The notion of a hard or difficult choice is ambiguous; it can mean that something is unlikely to be selected for a variety of reasons, or it can mean that the options engender moral conflict" (304–5). But the examples he gives to illustrate this ambiguity do not capture the sort of case we need here involving immediate background conditions that manipulate the likelihood of helping. He contrasts the scenarios of being tempted to accept a bribe because one has been socialized to value affluence with being tempted to accept it because it would pay for a life-saving operation for one's child. In one case, the moral responsibility of accepting the bribe is mitigated and in the other it is not. In my terms, *both* of these scenarios count as cases of normative overriding, in virtue of the fact that both the reason not to accept the bribe and the potentially mitigating reasons are consciously available.

might make this explicit by appeal to the distinction between an anti-altruistic motivation and a non-altruistic one. A person is *anti-altruistic* if she lets considerations that are not morally weighty carry more normative weight than helping another; for example, if she refuses to help—or merely avoids helping—because she may break a fingernail in doing so. In this case, altruism is undermined by reason-giving considerations. She takes doing something else to be more important than helping, or she lacks the will to help—the point is that she is actively responding to the reason to help another.[14] A person is *non-altruistic* simply when she is not behaving altruistically. By this I mean that she is not responding in any way to a reason to help another—she is not ignoring it, she is not rejecting it, she is not trying to respond or wishing that she could. Being altruistically motivated depends at least in part on not being anti-altruistic. It does not depend—at least, not straightforwardly—on not being non-altruistic. I suspect that if anything morally interesting is shown by these studies, it is the presence of this third alternative: one who is motivated neither by the suffering of others nor by one's own indifference or self-interest. The enemy of altruism is not merely egoism, but also being neither altruistic nor egoistic. It is implausible to think that the seminarians who did not stop hurried on because they were motivated by, say, apathy or self-interest or undue concern about their duty to give the volunteered talk on time. There's little reason to think they had any such normative reasons. We might call this *myopic motivation*. And yet this is not just because they were unreflective, but because, for

---

[14] Depending on how they're interpreted, the Milgram experiments might show that altruism—or at least compassionate behavior—is "normatively overridden" in this sense, by showing that certain reason-giving considerations diminish the frequency of helping behavior (Milgram 1963, 1974). On one interpretation of those studies, people's ordinary desires not to harm another are overridden by their motive to obey what they take to be legitimate authority figures. Given the robust anecdotal evidence of psychological stress observed (and also filmed) by the researchers, it would seem that here the motive to obey resides squarely in the realm of (subjective) reasons—albeit maybe not very good reasons. The subjects experience a value conflict that renders them too weak to do the right thing; their desire to help another is normatively overridden by their desire to follow instructions given by institutional experts.

However, since obedience is not obviously a morally trivial concern, the Milgram studies do not highlight situational features that we take to be morally *irrelevant* to the performance of the act of helping. On this interpretation, they involve substantive reasons-based conflict, in which one's reasons to obey apparent authority figures conflict with and override one's reason to be compassionate. (For an alternative interpretation, one that takes the Milgram studies to involve non-normative overriding, in which subjects act contrary to their dispositions to help because of a lack of confidence in their own judgment and a corresponding fear of embarrassment that is not reason-providing, see Sabini and Silver 2005.) The pertinent question is whether the subjective normative reasons are morally trivial ones. Though the Milgram studies *do* help to establish that altruism is undermined, they do not show that it is as easily undermined as would be needed to abandon philosophizing about the motive. To do that, we would need evidence that morally irrelevant features serve as subjective reasons that normatively override an agent's subjective reason to be altruistic.

Cases that demonstrate the bystander effect might seem to provide a better candidate for this role—the role of showing that morally irrelevant reason-giving considerations diminish the frequency of helping behavior, and thus that the motive of altruism is "normatively overridden"—to the extent that we take the relevant motive (say, fear of embarrassment) to be a normative reason that explains the subjects' inhibition to act. (For overviews of such cases, see Latané and Darley 1970, and Latané et al. 1981.) But it is hard to accept that this is the correct interpretation of such studies; the relevant situational factors at play in those cases do not provide subjective reasons in the required sense. Indeed, the bystander studies are interesting in that they seem, like the Good Samaritan study and possibly unlike the Milgram studies, to pick out a situational influence that does not qualify as a reason-giving consideration. Same with the other studies mentioned at the outset of this paper. The presence of others, like the presence of a dime or the smell of baking cookies or the fact that one has siblings, all seem to operate on a level of awareness that is not sufficiently conscious to make these considerations deliberatively relevant, or to bring them into the realm of reasons for action. Indeed, these most discussed studies in social psychology seem to show that subjects' desires to do the right thing are *non*-normatively inhibited by situational features; they do not involve value conflict, or show that the subjects have bad values.

all we know, they might not have seen the situation as a normative one—let alone a moral one; most probably didn't see that a question of 'should' had come up.

Let me forestall two possible misinterpretations. First, I do not claim that the subjects in the study are not blameworthy. For one thing, it's impossible to tell from the study itself whether some subjects were morally callous or acted for bad reasons. For another, some subjects may be blameworthy for failing to notice a person in need. I do not want to take a stand here on whether one can be morally responsible for such ignorance. One might think that there is a kind of absent-mindedness that is culpable, and one might think that one *should* be paying attention. That would be compatible with my position. The second clarification is this. My discussion of normative and non-normative overriding may appear to suggest an inaccurate picture of moral motivation. This is because we often do not engage in any sort of *weighing* process when motivated by altruism. And I do not want to imply that we do. So I do not want to yoke moral choice to a process of weighing or ranking options, though I do think that we can usefully appeal to such evaluative rankings in our normative explanations of behavior.

My central claim is that there seems to be a sense in which the normative thought of the subjects, "I ought not to be late", can provide them with a normative reason to go into the lecture hall, *without* its reflecting on their capacity for being altruistically motivated—even when noticing someone in trouble. The criterion, then, for what might count as an empirical threat to philosophical accounts of altruism is this. To cast doubt upon whether altruism plays a strong enough role in the explanation of behavior, one must show that morally trivial reasons in general prevent or undermine altruistic motives.[15]

In principle, it would be possible to provide an empirical design for settling whether a person is anti-altruistic. Given our practices of explaining and predicting behavior, we can and do develop plausible accounts of a person's motives and reasons in particular cases. It's true that there are substantial methodological difficulties in gathering such evidence, but these have not prevented psychologists in engaging productively with the issue. In fact, as Ross and Nisbett explain, it is significant that social psychology "was the one field of psychology that could never really be 'behaviorized'. Its most astute practitioners always understood that it is the situation as construed by the subject that is the true stimulus. This meant that theory was always going to have to focus on subjective interpretations of stimuli and responses as much as on stimulus–response relationships themselves" (1991, 11). So, while there is a longstanding debate among philosophers about how normative reasons can be empirically tractable, solving that metaethical issue is not necessary in order to study people's reasons. Though it may not be an easy empirical matter (simply asking people what they think won't work)—it is possible to do.[16] Daniel Batson for example, has investigated the role that empathy might play in prosocial behavior (for an overview, see his 1991, and also Sober and Wilson 1998). Keep in

---

[15] In a paper defending virtue ethics, Sreenivasan suggests that "the fact that one is in a hurry can defeat the reason to help someone in distress. Naturally, it depends upon a comparison of various factors, such as the importance of what one is hurrying for and the nature of the victim's distress" (2002, 60). Though I agree with his general claim, this particular passage misses the point. The very question of justifying reasons cannot come up until one can identify a subject's reasons. Indirectly, Sreenivasan seems to want to address this when he argues that the relevant studies must take into account the subjects' construal of their situation. But the particular test he proposes for ensuring this—that "the subject and the observer must agree" on whether a behavioral measure specifies a response that is paradigmatic of the relevant virtue, the reason for which is not defeated by reasons elicited by the situation (61–62)—does not go far enough. Whatever one's method for determining whether there is such agreement, it is reasonable to think that most of the subjects in this study *would* agree with the experimenters' (and our) normative judgments regarding their case. The criterion he proposes does not capture what reasons a subject has in the study conditions.

[16] This point is addressed in Sreenivasan (2002).

mind, too, that the difficulties involved in predicting behavior—having to do with the myriad differences among people's cognitive and motivational systems, and with the changeability of their internal states over time—do not in themselves cast doubt on the fact that people do have reasons for acting. Such reasons are, in fact, adduced in countless empirical studies offering explanations of behavior.

I've argued that the experimental data do not reveal whether altruistic motives are what I call normatively overridden, and that, in order to challenge the possibility/efficacy/strength of altruism, one must establish that altruism is (regularly, or often enough) normatively overridden. One cannot show that altruism is explanatorily useless by showing that non-reason-giving factors decrease the probability of acting altruistically. To do this, one would need to establish something that we already know is false; there would have to be a near-total prevalence of cases in which reason-giving-considerations override altruistic motives. This condition has been presented as a necessary one for rendering philosophical accounts irrelevant. The next section will challenge this claim.

4 The Challenge Pressed Further

The argument in Section 3 was driven by the simple observation that there are countless situations in which people fail to help others (you and I are presumably in such a situation now), and that very few of these are relevant to the question of whether the motive of altruism is a robust one. I argued that the question of which cases are relevant could be determined only once we know a subject's reasons for acting, and that the relevant studies do not identify those. In this section, I want to consider the objection that that argument may have incorrectly restricted the arena. I've mentioned that one of the lessons of the situationist evidence is that altruism is threatened not merely by indifference or self-interest, but by something that I call 'myopic motivation', in which helping others does not appear to be even remotely relevant to the question of what to do. I brought up the possibility that the question of relevance might be affected by what sort of knowledge and perceptions one is responsible for having in the first place. Shouldn't the subjects have noticed the need for help? This gets at the basic and intuitive idea that behaving altruistically is not merely a matter of helping another, but is a matter of doing so intelligently (that is, of correctly perceiving the situation, engaging in the proper sort of deliberation, and responding in the appropriate way).

After all, while it's true that knowing a person's reasons would help to determine whether they were altruistic, it seems equally plausible that if certain reasons never occur to one—if, say, helping another rarely seems like a relevant option, or seems relevant only because certain morally irrelevant features are present, then we can say that the motive of altruism is not robust, even without meeting the condition of normative overriding that I presented above. It would remain the case that situational features pervasively undermine the capacity for altruism.

In sum, the original challenge can be put forward once again at a higher-order level. The objection allows that it is both conceptually and empirically possible for altruism to occur under the right situational conditions—and can even grant that this happens often. As it should: When subjects *do* perceive unambiguously that another needs help (and they are not in a hurry, and there are few bystanders around, and so on) an overwhelming number help out.[17] The

---

[17] For example, the bystander effect disappears when the situation is unambiguous. As Doris observes "[n]umerous studies of staged emergencies have found impressive rates of intervention, in some conditions approaching 100% [....]. The situationist point is not that helping is rare, but that helping is situationally sensitive" (2000, 35. Cf. also p.19).

argument above, then, might prompt the following response. Altruism might indeed play a strong role *when* people act for reasons *and* are aware of the relevance of altruistic reasons for acting. That is, it might be correct that, if and when certain considerations are before one's mind, altruism wins out, at least often enough. Moreover, we might even grant that the experimental subjects in question can be said to behave only non-altruistically (and not anti-altruistically), and that the experimental data do not show that altruistic motives are normatively overridden. The objection goes on to say that if that is all I have established so far, then I have not yet shown that philosophical accounts of altruism are justified as empirical theses. I may have identified a sufficient (and, as it happens, non-instantiated) condition for undermining the role of altruism in these theories, but I have not provided a necessary condition for doing so. This is because if the very conditions for being normatively motivated (such as perceiving that someone needs help) are themselves situationally determined, then that would be a blow to those accounts—even though, again, it is true that people may have altruistic motives in the right circumstances. The idea is that the deeper-level situational explanation would undermine the relevancy of altruism as an explanatory state.

I have two, logically independent, replies to this. First, I see no reason why having a deeper-level situational explanation undermines the robustness of altruism itself as a motive worthy of philosophical investigation. One may still need to invoke altruism as such to explain the particulars of the action. Understanding how a subject sees a situation, and what her reasons are, is central to situationism itself, let alone our practices of attributing responsibility. After all, there's evidence that even determinism does not rule out normative explanations. The fact is that explanations invoked in attributions of moral responsibility are not restricted to merely causal explanations of external, behavioral, manifestations, but routinely draw upon an agent's evaluative attitudes and intentions.[18]

The second reply is simpler and more direct. It begins by acknowledging that one cannot ensure that one will not be influenced by situational features in ways that are at odds with one's values. As the many relevant studies in social psychology suggest, there are too many situational variables to be on the lookout for. Moreover, in general people are abysmal at predicting what sort of influence a situational feature will tend to have; so they won't have a clue what to be on the lookout for.[19] Not to mention that many will be unable to resist such influence even when they can see what is happening to them. Even so, it remains true that many people do recognize the need to control or mitigate the influence of their environment in order, say, to behave with decency toward others. The question is, what explains these attempts to mitigate situational influences? *Some* philosophical account of altruism will be worth developing to capture the nature of such a motive. In other words, people are capable of seeing the relevance of the very factors that initially seem irrelevant to the exercise of the motive. In doing so, we make room for normative guidance in our lives, and *that* leaves ample room for philosophers to have substantive debates about what this explanatory, normative feature might consist in. This reply appeals to the fact that different

---

[18] The assignment of responsibility based on the attribution of evaluative attitudes is a distinct enterprise from that of providing a causal explanation, and the two can come apart. For one interesting illustration of this, see the experiments in Woolfolk et al. (2006).

[19] This is, surprisingly, true even when people have been apprised of particular situational effects; they are in general very poor at generalizing what they've learned when making causal attributions in very similar cases. Indeed, in Pietromonaco and Nisbett (1982), it is noteworthy that subjects had difficulty accurately recalling the results of Darley and Batson, let alone generalizing the lessons learned from that study about the effect of hurrying. Moreover, as Pietromonaco and Nisbett report: "Fully 36% recalled, incorrectly, that the personality variable had affected helping [in that study]. Since the subject population and materials were similar to those used in previous studies, we must suspect that prior beliefs distorted the perception or recall of the information" (1982, 3).

levels of explanation are at play here. We can, and do, often enough attend to and overcome situational features that might undermine our altruistic instincts—and we can only do this precisely *because* we have altruistic motives. If altruism can explain attitudes toward (as well as actions aimed at controlling or mitigating) the influence of situational features, then it is robust enough to vindicate efforts to develop a philosophical account of this phenomenon.

To see how the two lines of reasoning in this section and the last one fit together, it might be helpful to frame the issues in terms of what altruism requires. One might think, simply, that altruism requires that one help out when one notices the opportunity[20] for help and there are no reasons that would justify doing something else instead. If this is what is required, then the argument of Section 3 is sufficient. The Good Samaritan study has not established that this condition obtained for the subjects; thus it could not show that 'being in a hurry' is a candidate for something that might count as a justifying reason for not helping. Relevant to the opportunity to help another, it doesn't count as a reason at all—not because of its lack of importance, but because of its role in the subject's mental economy. Suppose the subjects were hurrying not to give a speech, but to offer real help to a group of people—say, to perform life-saving surgery. It would be tempting then to say that they did not stop because they did not want to jeopardize or delay their good deeds. But again, even this would be an unwarranted interpretation. There would be no reason, given the experimental methodology discussed above, to think that they were necessarily motivated by such a justifying reason. Or rather, one might agree that they had a reason for hurrying on, but insist that it is an open question whether they had a reason for failing to help the distressed person by the wayside. As I've said above, whatever your philosophical account of altruism, the motive of altruism is nothing if not a response to a reason to help others. As such—as a response to this sort of reason—altruism exists in the normative realm of practical reasons. Its existence cannot be undermined by evidence that non-reason-giving factors diminish the frequency of helping behavior.

On the other hand, to repeat, one might think that altruism requires more than that. One might think that it requires that one notice or recognize certain things, such as that people are in need and are worthy of being helped in virtue of that need (and requires, moreover, that one is thereby properly motivated to act). Say it turns out to be the case that situational features pervasively undermine the possibility of having altruistic reasons to act. The condition I offered—concerning non-normative overriding—in Section 3 would be satisfied. Yet it wouldn't be right to say that altruism has been thereby secured as an empirically worthy topic of investigation. The prevalence of non-reason-giving factors could undermine the robustness of altruism as it is instantiated in the actual world. But I've argued in this section that, as a matter of fact, it doesn't. For the skeptical conclusion about philosophical accounts of altruism to be justified, we would have to show not only that such a state of being non-altruistic is pervasive, but also that it is something that most people could—and would—do nothing about. This point, that we can improve our moral behavior by becoming aware of the potential for situational influences, is often made by situationists themselves.[21] Moreover, the very fact that we are interested in the possibility of morally improving ourselves by learning of the effects of situational features of the environment

---

[20] What counts as an 'opportunity', and how pressing it must be to constitute a moral requirement, is of course a controversial issue in moral philosophy. For one compelling account based on the notion of physical proximity, and relevant to the present study, see Jeremy Waldron (2003).

[21] Cf. especially Doris (1998) and Harman (1999) for thoughts on some of the moral effects and benefits that can follow from appreciating situationism and the fundamental attribution error.

itself points to the existence of a strong altruistic motive: a normative motive, strong enough to warrant philosophical investigation.

References

Batson CD (1991) The altruism question: toward a social-psychological answer. Erlbaum, Hillsdale

Campbell J (1999) Can philosophical accounts of altruism accommodate experimental data on helping behaviour? Australas J Philos 77(1):26–45

Darley JM, Batson CD (1973) "From Jerusalem to Jericho": a study of situational and dispositional variables in helping behavior. J Personal Soc Psychol 27(1):100–108

Doris JM (1998) Persons, situations, and virtue ethics. Nous 32(4):504–530

Doris JM (2002) Lack of character. Cambridge University Press, New York

Harman G (1999) Moral philosophy meets social psychology: virtue ethics and the fundamental attribution error. Proc Aristot Soc 99(3):315–331

Jamieson D (2002) Sober and Wilson on psychological altruism. Philos Phenomenol Res 65(3):702–710

Latané B, Darley J (1970) The unresponsive bystander: why doesn't he help? Appleton, New York

Latané B, Nida SA, Wilson DW (1981) The effects of group size on helping behavior. In: Rushton JP, Sorrentino RM (eds) Altruism and helping behavior. Erlbaum, Hillsdale

McDowell J (1979) Virtue and reason. Monist 62:331–350

Milgram S (1963) Behavioral study of obedience. J Abnorm Soc Psychol 67:371–378

Milgram S (1974) Obedience to authority. Harper & Row, New York

Mischel W (1968) Personality and assessment. John Wiley, New York

Nagel T (1970) The possibility of altruism. Oxford University Press, London

Pietromonaco PR, Nisbett RE (1982) Swimming upstream against the fundamental attribution error: subjects' weak generalizations from the Darley and Batson study. Soc Behav Pers 10(1):1–4

Ross L, Nisbett RE (1991) The person and the situation: perspectives of social psychology. McGraw-Hill, Boston

Sabini J, Silver M (2005) Lack of character? Situationism critiqued. Ethics 115:535–562

Schoeman F (1987) Statistical norms and moral attributions. In: Schoeman F (ed) Responsibility, character, and the emotions. New York, Cambridge

Slote M (2001) Morals from motives. Oxford University Press, Oxford

Sober E, Wilson DS (1998) Unto others: the evolution and psychology of unselfish behavior. Harvard University Press, Cambridge

Sreenivasan G (2002) Errors about errors: virtue theory and trait attribution. Mind 111(441):47–68

Waldron J (2003) Who is my neighbor?: humanity and proximity. Monist 86(3):333–354

Woolfolk RL, Doris JM, Darley JM (2006) Identification, situational constraint, and social cognition: studies in the attribution of moral responsibility. Cognition 100(2):283–301