

Egoism, Empathy, and Self-Other Merging

Joshua May (Monash University)

Penultimate draft; final version in:

Southern Journal of Philosophy [Vol. 49 \(2011\)](#), pp. 25-39

Spindel Supplement: Empathy & Ethics, Remy Debes (ed.)

Emerging Scholar Prize Essay

Abstract: Some philosophers and psychologists have evaluated psychological egoism against recent experimental work in social psychology. Dan Batson (1991; forthcoming), in particular, argues that empathy tends to induce genuinely altruistic motives in humans. However, some argue that there are egoistic explanations of the data which remain unscathed. I focus here on some recent criticisms based on the idea of self-other merging or “oneness,” primarily leveled by Robert Cialdini and his collaborators (1997). These authors argue that the putatively altruistic subjects are acting on ultimately egoistic motives because empathic feelings for someone in distress tend to cause them to blur the distinction between themselves and the other. Employing a conceptual framework for the debate, I argue that the self-other merging explanation fails to explain the empathy-helping relationship on primarily non-empirical grounds, regardless of the empirical results Cialdini and colleagues report.

1. Introduction

Psychological egoism is roughly the thesis that our actions are always ultimately motivated by self-interest. Philosophers don't have much sympathy for this view. Many appear to think it subject to a powerful dilemma: either it is the trivially true, yet uninteresting, claim that all of one's actions are motivated by one's own motives or else it is the controversial thesis it claims to be, in which case it is obviously false (Feinberg, 1978, p. 496).¹ Clearly the view is supposed to be the substantive thesis described on the second horn. Yet it's unclear why one should consider it obviously untrue. The consensus among psychologists (and common sense) is that a great number of our mental states, even our motives, are not accessible to consciousness. While introspection isn't a worthless source of knowledge of our own minds, it is quite limited in its ability to discern the nature of our ultimate motives that very well may be subconscious. For this and several other reasons, it is rather surprising that

¹ Cf. Hume (1751/1998, App. 2.9, p. 167), Hutcheson (1725/1991, esp. p. 278, Raphael sect. 327), and Bernard Williams (1973, pp. 262-3). They all consider various examples of actions they deem rather implausible to construe as ultimately motivated by self-interest.

psychological egoism is so widely and immediately rejected in the philosophical community.²

More recently, some philosophers, such as Elliott Sober, have evaluated psychological egoism against experimental work in social psychology, primarily done by C. Daniel Batson. Batson (1991; forthcoming) argues that empathy in particular tends to induce genuinely altruistic motives in humans. While he admits that more studies can and should be done, Batson maintains that the experiments conducted on all the various egoistic hypotheses show that we are at least tentatively justified in believing what he calls “the empathy-altruism hypothesis,” which entails psychological egoism is false. Batson marshals a non-deductive argument for his conclusion that “Contrary to the beliefs of Hobbes, La Rochefoucauld, Mandeville, and virtually all psychologists, altruistic concern for the welfare of others is within the human repertoire” (1991, p. 174).

However, some philosophers, psychologists, and biologists argue that Batson’s work fails to rule out all the relevant egoistic explanations of the data. While there are a range of opponents, I want to focus here on some recent criticisms based on the idea of self-other merging or “oneness,” primarily leveled by Robert Cialdini and his collaborators. These authors argue that the putatively altruistic subjects are acting on ultimately egoistic motives because empathic feelings for someone in distress tend to cause them to blur the distinction between themselves and the other.³ Employing a conceptual framework for the debate I have developed previously, I argue that the self-other merging explanation fails to explain the empathy-helping relationship on conceptual grounds, regardless of the experiments Cialdini et al. (1997) report.

2. Understanding Egoism

While Batson has a keen understanding of the egoism-altruism debate, I have attempted to develop a richer framework (see May, forthcoming), which builds on the work of Elliott Sober and David Sloan Wilson (1998) in their extraordinary book *Unto Others*. Following Alejandro Rosas (2002, p. 98), we can begin by defining psychological egoism (and its counterpart, psychological altruism) most broadly and uncontroversially as follows. *Psychological egoism* is the claim that all our motivational states are ultimately egoistic; *psychological altruism* is the claim that some of our motivational states are ultimately altruistic.

² Michael Slote (1964) is one clear exception here to both trends: he takes psychological egoism as a serious empirical claim and argues that it may well be true. However, Slote appeals to fairly dated, behavioristic psychology.

³ This sort of egoistic merging objection is also anticipated (and rejected) by Hutcheson. He imagines someone explaining away our apparently altruistic desires concerning those who are dear to us: “Children are not only made of our bodies, but resemble us in body and mind; they are rational agents as we are, and we only love our own likeness in them” (1725/1991, p. 279, Raphael sect. 327). (Many thanks to Stephen Darwall and Tamar Schapiro for bringing this wonderful parallel to my attention.)

But what makes a state “egoistic” or “altruistic”? By all accounts, in both cases the motivational state or desire (in the broad sense) must first be *ultimate* (or what some might call “intrinsic,” or “non-instrumental”)—i.e. a desire for something for its own sake. Second, an altruistic desire must concern the *benefit* of someone other than oneself. However, contrary to Sober and Wilson, the desire mustn’t simply be “self-directed” in the sense of being about oneself or one’s own benefit (May forthcoming).⁴ What I have called the “essential beneficiary” represented in the content of the desire must be someone represented as distinct from oneself. I defined an essential beneficiary in the following way:

A subject S^* is an *essential beneficiary* of S ’s desire that p if and only if: it is a necessary condition for the satisfaction of S ’s desire that S^* benefits.
(p. 7)

One’s desires are all *egoistic* if the contents of one’s ultimate desires all fail to contain someone other than oneself as an essential beneficiary. To consider a concrete case, suppose Samantha lends Ben some money because she wants him to be happy. Furthermore, she doesn’t want him to be happy in order to make herself happy, or in order to avoid social sanctions, or in order to achieve any other purpose. Instead, Sam’s desire for Ben’s benefit is ultimate: she desires that state of affairs for its own sake. If even one of us is like Sam, then psychological egoism is false.

This conception of psychological egoism is of course only a more detailed version of the view that I mentioned philosophers tend to regard as easily refutable. Indeed, the only major philosophers who seem to have endorsed the view explicitly are Hobbes and Bentham. My task here is to object to an argument for egoism, so it had better be worth doing so. I previously mentioned arguments against egoism that appeal to introspection and common sense. While I suggested these are rather weak objections to the view, there are others that have received more attention. For example, many take Joseph Butler (1827) to have refuted egoism in his *Fifteen Sermons* (e.g. Broad 1930; Feinberg 1965, p. 497).⁵ Yet Sober and Wilson (1998, ch. 9) have shown that this argument fails to refute psychological hedonism. And the problems they raise apply equally to Butler-style arguments against psychological egoism generally. In fact, to my mind, Sober and Wilson show that all of the traditional philosophical arguments against the view are utterly lacking, including ones that appeal to triviality, falsifiability, Nozick’s experience machine, and the paradox of hedonism. So I submit that we likewise take psychological egoism as a live empirical claim. While we should not rule out the possibility that non-empirical, philosophical arguments can advance or even settle the issue, there is a body of empirical research that deserves examination.

⁴ Though I hadn’t mention it previously, this notion of a self-directed desire seems equivalent to what Bernard Williams calls “I-desires” (1973, p. 261). (Thanks to Richard Holton for bringing this to light.)

⁵ Broad proclaims that Butler “killed the theory so thoroughly that he sometimes seems to the modern reader to be flogging dead horses” (1930, p. 55).

3. Empirical Evidence for Altruism

Psychological egoism makes a claim about the motives people do in fact have (as opposed to those which they should have). So, in principle, empirical evidence may be gathered to bear on the issue. In fact, this is precisely what Batson and other psychologists have done in some of their work on empathy-induced helping behavior. Batson (1991), in particular, argues that the experiments conducted thus far provide evidence for an altruistic model he calls the *empathy-altruism hypothesis*, which holds that empathy tends to induce in us ultimate desires for the well-being of someone other than ourselves. Batson and Shaw (1991) frame the hypothesis as holding that “empathy evokes altruistic motivation directed toward the ultimate goal of reducing the needy person’s suffering” (p. 114).

Batson begins by concentrating on a robust effect of empathy on helping behavior discovered in the 1970s.⁶ The *empathy-helping relationship* is the finding that the experience of relatively high empathy for another perceived to be in need causes people to help the other more than relatively low empathy. However, as Batson recognizes, this doesn’t establish psychological altruism or egoism, because it only specifies behavior, not the ultimate motive, much less whether it is altruistic or egoistic. Given that there can be both egoistic and altruistic explanations of the empathy-helping relationship, Batson and others have devised experiments to test some of their differing predictions.

The general experimental approach involves placing human subjects in situations in which they have an opportunity to help someone they believe is in need while manipulating other variables in the situation. The idea is to provide circumstances in which egoistic versus altruistic explanations of empathically induced helping behavior make different predictions about what people will do. Different hypotheses provide either egoistic or altruistic explanations of why the subjects ultimately chose to help, depending on whether it posits an egoistic or altruistic ultimate desire. Of course, this requires several familiar assumptions about human action, motivation, and practical reasoning in order to infer motives from behavior, but they are fairly uncontroversial.⁷

⁶ Empathy here, according to Batson (1991), is roughly “an other-oriented emotional reaction to seeing someone suffer” (p. 58) that “includes feeling sympathetic, compassionate, warm, softhearted, tender, and the like” (p. 86). This is to be distinguished primarily from “personal distress,” which is “made up of more self-oriented feelings such as upset, alarm, anxiety, and distress” (p. 117). This is somewhat stipulative, but it’s unproblematic, so long as it captures the rough idea and we keep the same notion fixed throughout.

⁷ In short, we merely need to assume that people have mental states with intentional content that play some causal role in bringing about action. And such mental states need not be immediately available to the agent’s consciousness at the time of action or otherwise. For a more detailed discussion of the background assumptions involved, see Batson (1991, pp. 64-67), Sober and Wilson (1998, Ch. 6), and Stich, Doris, and Roedder (2010).

Several egoistic explanations of the empathy-helping relationship are in competition with the empathy-altruism hypothesis. Each one claims that experiences of relatively high empathy (“empathic arousal”) cause subjects to help simply because it induces an egoistic ultimate desire; the desire to help the other is solely instrumental to the ultimate desire to benefit oneself. However, the experiments seem to rule out all the plausible (and some implausible) egoistic explanations. Generally speaking, the data seem to show that the empathy-helping relationship is not put in place by egoistic ultimate desires to either: (a) relieve personal distress or aversive arousal, (b) avoid self-punishment (e.g. feelings of guilt), (c) avoid social-punishment (e.g. looking bad to others), (d) obtain rewards from self or others (e.g. praise, pride), (e) obtain feelings of vicarious joy at relief, or (f) gain a mood-enhancing experience (Batson 1991; forthcoming).⁸ Furthermore, according to Batson, the data all conform to the empathy-altruism hypothesis, which claims that empathic arousal induces in subjects an ultimate desire with something like the following content: the person in need is helped.

While Batson admits that more studies can and should be done on this topic, he ultimately concludes that we are at least tentatively justified in believing that the empathy-altruism hypothesis is true. He marshals a non-deductive argument for this conclusion:

In study after study, with no clear exceptions, we find results conforming to the pattern predicted by the empathy-altruism hypothesis.... At present, there is no egoistic explanation for the results of these studies.... Pending new evidence or a plausible new egoistic explanation for the existing evidence, the empathy-altruism hypothesis, however improbable, seems to be true. (1991, p. 174)

Thus, Batson maintains that we are justified in believing that psychological egoism is false. At this point, the psychological egoist can only explain the data if she can provide a similar, yet still egoistic, hypothesis.

4. Self-Other Merging

Many attempts to explain the empathy-helping relationship in egoistic terms have apparently failed. However, more recently, some have leveled a new kind of challenge to the empathy-altruism hypothesis. Empathy, one might argue, involves taking the perspective of another—walking a mile in her shoes, as it were—and in some sense feeling the pain of those with whom we empathize. Why not think then

⁸ Batson & Shaw (1991) provide a more condensed review of the empirical evidence for the failure of egoistic explanations of the sort enshrined in (a)-(f).

that, in helping (or offering to help) another, empathically aroused individuals are in some sense helping themselves, or at least have some such mental representation?⁹

As Batson, Sager, Garst, Kang, Rubchinsky, and Dawson (1997) point out, the materials for this kind of response may be found in the work of Daniel Wegner, Harvey Hornstein, and Art Aron, among others. And the basic idea has been defended by Kristen R. Monroe, Michael C. Barton & Ute Klingemann (1990, esp. 122), but primarily on the basis of case studies and not in response to Batson's project. But the challenge is most explicitly formulated and directed against Batson by Robert Cialdini and his collaborators. Cialdini, Brown, Lewis, Luce, and Neuberg (1997) contend that empathic arousal blurs the distinction between self and other in such a way that the seemingly altruistic motives of these subjects are really non-altruistic: one is motivated to help the other for whom empathy is felt, but the "other" is represented in the content of one's ultimate desire as, to at least some extent, oneself. Empathically aroused individuals allegedly have an "experience of oneness—a sense of shared, merged, or interconnected personal identities" (p. 483). Cialdini and his collaborators write: "If the circumstances specified in the [empathy-altruism] model as leading to truly altruistic acts (interpersonal closeness and perspective taking instructions) are the same circumstances that enhance the merging of self and other... then one can doubt whether those helpful acts reflect the selflessness required for true altruism" (1997, p. 490).¹⁰

Cialdini et al. only claim to have shown that the empathy-induced motives are "nonaltruistic" as opposed to egoistic, because when "the distinction between self and other is undermined, the traditional dichotomy between selfishness and selflessness loses its meaning" (pp. 490-1). The idea seems to be that certain motives, while not egoistic can still be non-altruistic. While psychological egoism certainly claims that all of our motives are egoistic, any motive which is non-egoistic is incompatible with psychological egoism, even if it is not altruistic. An ultimate desire for the number of stars to be even, for example, concerns no one's benefit, not even one's own. The existence of such ultimate desires provides no support for psychological altruism, but they are nevertheless also incompatible with psychological egoism.¹¹ But Cialdini et al. do not seem to be positing the existence of these kinds of non-altruistic motives. Their oneness motives, assuming they make

⁹ Another recent and important challenge to Batson's project objects to his strategy of ruling out egoistic explanations *one at a time*. This issue is raised rather quickly and briefly by several authors, including Cialdini, Brown, Lewis, Luce, and Neuberg (1997, p. 483). But see Timothy Oakberg (ms) for a much more detailed and powerful defense of the challenge, though I am unsure whether it is successful (as are Stich et al. 2010, p. 201).

¹⁰ See also Neuberg, Cialdini, Brown, Luce, Sagarin, & Lewis (1997) and Maner, Luce, Neuberg, Cialdini, Brown, & Sagarin (2002).

¹¹ For further discussion of the relationship between egoism and altruism, see May (forthcoming) and Stich et al. (2010). An ultimate desire to do one's duty might also be a candidate for a motive that is neither egoistic nor altruistic. Unfortunately, it appears this possibility has received little attention in empirical research on empathy-induced helping behavior. (Thanks to John Deigh and Walter Sinnott-Armstrong for discussion on this point.)

reference to someone or other, make reference to both oneself and another. But it is not obvious that we should count such relational desires as non-altruistic (see the Tennis Game case in May, forthcoming). Moreover, Cialdini and his collaborators seem to be suggesting that their oneness motives are compatible with psychological egoism in addition to posing a challenge to the empathy-altruism hypothesis. In any event, we can examine whether their proposal supports psychological egoism in addition to whether it merely poses a challenge to psychological altruism. As we shall see, this is a fine line to walk for the empathy-merging hypothesis.

The empirical evidence Cialdini et al. (1997) advance in favor of the empathy-merging model involves teasing apart empathy from perceived oneness as variables and testing how predictive they are of helping. The crucial prediction of the empathy-merging hypothesis is that “empathic concern would no longer be predictive of helping after the influence of oneness was extracted but that oneness would be significantly predictive when empathic concern was partialled out” (p. 483). They purport to have produced data confirming such a prediction. Batson et al. (1997) address this challenge empirically by raising worries about confounds in the experiments of Cialdini et al. and reporting data from experiments without such problems that are commensurate with the empathy-altruism hypothesis (see also Batson, forthcoming). However, I want to raise a more fundamental and conceptual problem with the self-other merging challenge or the “empathy-merging hypothesis” as Batson et al. call it. In effect, though they do not always make this clear, what Cialdini and his associates are suggesting is that the “empathy-helping relationship” is really a *oneness-helping relationship*: perceived oneness is the real cause of increased helping behavior, and empathy is a mere “concomitant,” as they put it (p. 483). This claim, however, is insufficient for a non-altruistic explanation. They must supplement it with the further claim that an ultimate desire due to perceived oneness is not altruistic. That is the claim I primarily wish to assess.

5. Proposals and Problems

First, it is unclear that removing “oneness” from empathy is conceptually possible. Cialdini et al. want to pry these two apart, but one natural conception of “oneness” seems rather essential to the perspective-taking that is crucial to empathy. If so, then this would undermine their claim that it is oneness and not empathy that increases helping. Nevertheless, what is more instructive is to explore the details of this empathy-merging hypothesis to see how exactly it is meant to explain the empathy-helping relationship (or the oneness-helping relationship).¹² We must clarify what such metaphorical talk as “blurring the distinction between self and other” means. This requires specifying the content of the ultimate desire to determine whether it is

¹² I will continue to speak only of the empathy-helping relationship since the coming problems will apply just as well even if there is only a oneness-helping relationship. (Thanks to Jesse Prinz for discussion here.)

egoistic or altruistic, and then provide the proposed beliefs this connects with to provide an explanation of the empathy-helping relationship. We can explore the prospects of several versions of the self-other merging proposal, all of which face serious problems.

5.1 Proposal 1: Peculiar Beliefs

The simplest version of the empathy-merging hypothesis would just attribute the belief to empathically aroused subjects that they exist in two separate bodies. Their ultimate desire would be to benefit themselves, but they believe they exist in the other body. So they instrumentally desire to help themselves in their other incarnation, so to speak. This appropriately explains why they help and in a clearly egoistic way.

But this, I submit, is a wildly implausible belief to attribute to normal people, even if the belief is subconscious. We might even count it as delusional. Of course, on certain metaphysical views we are all one, in which case the belief that one exists in two separate bodies would not be an error, and so perhaps not a delusion. But even if such a view about the world is correct, we need an argument to show that people would believe it. After all, typical Western subjects (as these are) do not seem to hold such a metaphysical view. Moreover, it is even more implausible to attribute the relevant belief given that it must be *empathy-specific*. That is, to explain the empathy-helping relationship, one must explain the helping behavior by appealing to some feature of empathy or something that empathy causes, since only the empathically aroused subjects help in such higher numbers. So not only must we attribute what appears to be a delusional belief to subjects on the self-other merging account, we must also hold that empathy in particular induces such a belief.

It is important here to keep in mind that we needn't rule out every egoistic explanation, however implausible it may be. We don't, for example, seriously consider any hypotheses involving aliens from outer space playing some role. In a similar context to ours, Dale Jamieson makes something like this general, theoretical point: "The tortured, *ad hoc* explanations offered by [psychological egoism] are more like those given by the Ptolemaic view than those provided by relativity theory or quantum mechanics. All things considered, it is more plausible to suppose that [psychological altruism] is true than [psychological egoism]. The experiments of Batson and others are part of what supports this judgment" (2002, p. 706). The general point applies in this context as well. In any event, Cialdini et al. (1997) do themselves attempt to avoid attributing to subjects such peculiar beliefs: "What is merged is conceptual, not physical. We are not suggesting that individuals with overlapping identities confuse their physical beings or situations with those of the other" (p. 482).

5.2 Proposal 2: Indeterminate Identities

Cialdini and his collaborators only maintain that the representations of oneself versus another are to some extent “blurred.” One way to read this is that the identities of the people represented in the content of the relevant mental states are indeterminate.¹³ One problem with this proposal, though, is that it requires a substantial claim about the psychological ability to represent indeterminate identities of persons. Cialdini et al. do mention previous research they claim suggests the existence of such an ability—the “contemporary theory of the self” as dynamic and malleable. They also mention the idea of “inclusive fitness” in evolutionary theory, which they take to suggest that we can see ourselves in others by evidence of shared genes (pp. 482-3). But this is all quite compatible with explanations employing less drastic conceptual resources. The psychological research merely reveals phenomena such as “more similar self-other word associations” (p. 482); and evolutionary theory only shows that we are attune to the presence of similar genes and so forth. Such data do not preclude the idea that we are still only representing determinate identities of self and other and connecting them in various ways.¹⁴ In any event, even granting the ability to represent indeterminate personal identities, there are two additional problems with using this as a non-altruistic explanation of the empathy-helping relationship.

First, it won’t adequately explain the behavior of such empathically aroused individuals. The more fundamental purpose of attributing these ultimate motives to people is to explain and predict their behavior, or at least to develop an account that can. Other egoistic models purport to explain the helping behavior by holding that the motivation to help the other individual is simply instrumental to the ultimate motivation to help or benefit oneself. For example, one egoistic explanation (“aversive-arousal reduction”) holds that helping is merely a means to eliminating empathic arousal, which is hypothesized to be a feeling to which subjects are quite averse. This explains the behavior by positing a desire, here an instrumental one, to help the other, where this is represented wholly as an individual distinct from oneself. This is then coupled with the ultimate desire to benefit oneself and a belief (among others) that one can do this by helping the other when no one else apparently can. These materials are sufficient to explain, in an egoistic way, why the

¹³ Another quite different proposal would be that the content of one’s ultimate desire makes reference to oneself and the other, conceived as determinate and distinct. These are what I have called “relational desires,” and their employment here would amount to abandoning the crucial idea of merging. See May (forthcoming) for a detailed discussion of how relational desires figure into the egoism-altruism debate.

¹⁴ In a somewhat different context—discussing studies of rescuers of Jews in Nazi Europe—Neera Badhwar (1993, §2) briefly raises some related worries. Against Monroe et al. (1990), she cautions against jumping too quickly to attributing to subjects a loss of perceived individual identity based on self-reported feelings of oneness with another or a group.

empathically aroused individual performs the action of helping a person physically distinct from oneself.¹⁵

What, on the empathy-merging model, is going to likewise explain why empathically aroused subjects tended to help someone distinct from themselves?¹⁶ Recall that Cialdini et al. maintain that what is merged is “conceptual, not physical.” However, the two cannot be pried apart so easily. A conceptual representation is meant to guide behavior to the physical counterpart. My desire for some coffee, for example, must link at least with a belief that there is coffee in front of me in order to explain my picking it up and drinking it. What is crucial in this toy example is that the explanation of my action of interacting with an object at a distance from myself must be grounded in mental representations of myself as distinct from the object that is at a distance from me. Similarly, a conceptual representation of oneself as *not* at a distance from oneself will not—barring delusions, etc.—appropriately explain one’s behavior of interacting with objects that are distinct from oneself.¹⁷

It seems, then, that we need a representation of another as distinct from oneself to avoid delusional beliefs and to provide an explanation of the helping of a distinct individual. But where exactly does this enter into the mental states of the individual? This representation of distinct identities could occur in the belief alone. But then what is the ultimate desire? If empathically aroused subjects believe that the other person is distinct from themselves, the ultimate desire cannot be to help another—that would be altruistic. The only egoistic ultimate desire will be to connect helping the other person with benefit to oneself. But this is the model for egoistic explanations past, not the newfangled empathy-merging explanation, which appeals to the merging of personal identities to secure a non-altruistic explanation of the empathy-helping relationship. I have argued that this cannot be done in a way that appropriately explains the behavior of helping a distinct other. But let us set aside this problem and raise another.

Even if one could explain the behavior by appealing to indeterminate identities, the second problem is that this does not necessarily amount to a non-altruistic motive. If one acts on a motive that represents another’s benefit at all, even partially, then we have the materials for at least a partially altruistic motive. This relies to some extent on the conceptual framework I have previously developed,

¹⁵ Not in all the experiments did the subjects eventually get the chance to provide the help since they were usually debriefed prior to the opportunity (and so were apprised of the fact that there was no one in real need). But the same points apply to explanations of the act of offering to help and even to being so prepared.

¹⁶ Timothy Oakberg (ms) also raises this question, albeit briefly, against Cialdini and his fellow researchers. He doesn’t explore the potential answers and their problems, though.

¹⁷ Batson & Shaw (1991, p. 160) seem to interpret views like the empathy-merging hypothesis as *eliminating* the distinction between self and other such that the motive is neither egoistic nor altruistic. Batson et al. (1997, p. 497) appear to later endorse this construal as well. They are right that the *elimination* of self or other would yield this result and so pose a threat to altruism (and egoism). But we can now see that this is a rather implausible version of the empathy-merging hypothesis, as it will fail to appropriately explain the helping of individuals at all (whether self or other).

which I haven't argued for here (May forthcoming). But we can provide some general, additional grounds for such a conclusion. In short, admitting partial representations of another in the content of one's ultimate desire is a significant concession on the part of psychological egoists. After all, their view is not meant to be ultimately confined to empathy-induced helping. They are making a claim about the nature of *all* ultimate desires of humans—namely, that they are egoistic. If we allow the existence of partially altruistic motives, then the project of a parsimonious account of human motivation is undermined. Once the gates are open to a pluralism of pro-social motives, there is little reason to stop at admitting only partially altruistic ones. The theoretical case for psychological egoism employing principles of parsimony could not be maintained if the competing views both posit two distinct kinds of motives. At best we would be left with no theoretical grounds for preferring the empathy-merging to the empathy-altruism hypothesis. In fact, at such a point we may have grounds for preferring the latter given its likely support from introspection: doesn't it seem, for example, that your motivation to promote the well-being of your children, say, isn't instrumental to any other desire to benefit yourself? While introspection's value may be suspect in gathering the primary data, we might reasonably use it as a sort of tie-breaker.

It seems, then, that the self-other merging explanation is subject to a formidable dilemma. In order to properly explain the behavior of helping *another*, it must posit a representation of another. But if a partial representation of another in the content of the ultimate desire can explain such behavior, then it is arguably altruistic. The alternative is to attribute to empathically aroused subjects the belief that they exist in two separate bodies. But this is exceedingly implausible; we have no reason to believe that empathy induces such peculiar beliefs. So far, then, at best the empathy-merging explanation is either wildly *implausible* or can't *explain* the helping behavior and is potentially *altruistic*. I see only one final way of interpreting the merging explanation.

5.3 Proposal 3: Properties

At this point, the challenger might attempt to appeal to merging representations of *aspects* or properties of self and other. Cialdini and colleagues do at times seem to conceive of their proposal along such lines when they suggest that empathically aroused people “locate more of themselves in the others to whom they are closely attached” (p. 483). Here the explanation of the act of helping another is supposed to be explained, it seems, by appealing to an ultimate desire to benefit oneself combined with a belief that aspects of oneself are, as it were, over there in that other body. This provides the tools for explaining the behavior of helping another distinct from themselves since two individuals are represented—it is only the properties that are identified.

But this has two problems, which aren't mutually exclusive. First, it seems again rather implausible to attribute such beliefs to subjects, especially when the

attribution must be empathy-specific. Furthermore, even if empathically aroused subjects do believe that aspects of themselves exist in other bodies, we are still left unable to appeal to *non-altruistic* desires while still explaining the behavior. Ultimately, these “aspects” must be properties of individuals, and believing that one *shares* certain properties with another doesn’t entail that the ultimate desire of the individual is not to help another. Even believing that one’s *own* properties are instantiated in another does not clearly support an egoistic motive. Whether it is egoistic depends on the property. Perhaps, for example, the motive would be egoistic if the self-other property identity is specifically the property of being happy (or some such benefit). For example, one might hold that the ultimate desire of the agent is to benefit herself and that she believes she can do this by promoting the property of being happy (e.g.) in this other person, since she also believes this is the exact same property of herself. After all, empathy is often induced by perceiving the other to be similar to oneself.¹⁸

However, this explanation would then be too narrow to account for the empathy-helping relationship in its entirety. Research on empathy indicates that there are various similarities (e.g. same gender, age, background, etc.) that can induce higher levels of empathy for another perceived to be in need. People need not believe they are similar to the needy other in respect of harm or benefit to demonstrate the empathy-helping relationship (see the discussion of the “similarity manipulation” of empathy by Batson, 1991, p. 114). While it is plausible that normal subjects will always believe that the other will benefit from help, there is no reason to think that empathy always induces the belief that they are similar in respect of that benefit or harm, let alone identical.

6. Conclusion

Once we take psychological egoism seriously as a live thesis about our ultimate motives, there is every reason to address empirical work that bears on it. While Batson and his collaborators seem to have provided a great deal of support for the empathy-altruism hypothesis, it is quite natural to suspect that the role of perspective-taking and similarity in manipulating empathy is more amenable to an egoistic explanation of the empathy-helping relationship. However, as we’ve seen, the details of such a self-other merging explanation lead to deleterious results. On its most plausible construal, the empathy-merging hypothesis cannot appropriately explain the increased helping behavior of empathically aroused individuals, which is the name of the game in this approach to the egoism-altruism debate. While the considerations raised in this paper certainly don’t rule out all egoistic explanations of

¹⁸ There is some evidence that empathically aroused individuals do not ultimately desire to be the ones in particular to help (Batson 1991, pp. 149-53). This is compatible with the egoistic explanation currently being considered. It need only attribute to subjects the belief that they benefit even if they aren’t the ones in particular to help; their property of being happy that is instantiated in the other person can be promoted by anyone.

the empathy-helping relationship, they fend off one that is especially plausible on its face.¹⁹

References

- Badhwar, Neera (1993). "Altruism Versus Self-Interest: Sometimes a False Dichotomy." *Social Philosophy and Policy* 10 (1):90-117.
- Batson, C. D. (1991). *The Altruism Question: Toward a Social-Psychological Answer*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Batson, C. D. (forthcoming). *Altruism in Humans*. Oxford University Press.
- Batson, C. D., K. Sager, E. Garst, M. Kang, K. Rubchinsky, and K. Dawson (1997). "Is Empathy-Induced Helping Due to Self-Other Merging?" *Journal of Personality and Social Psychology* 73(3): 495-509.
- Batson, C. D. & Laura L. Shaw (1991). "Evidence for Altruism: Toward a Pluralism of Prosocial Motives." *Psychological Inquiry* 2 (2): 107-122.
- Broad, C. D. (1930). *Five Types of Ethical Theory*. London: Routledge, 2000. (Originally published in 1930 by Kegan Paul, Trench, Trubner & Co. Ltd.)
- Cialdini, R. B., S. L. Brown, B. P. Lewis, C. Luce, & S. L. Neuberg (1997). "Reinterpreting the Empathy-Altruism Relationship: When One Into One Equals Oneness." *Journal of Personality and Social Psychology* 73(3): 481-494.
- Feinberg, Joel (1978/1999). "Psychological Egoism," In Russ Shafer-Landau & Joel Feinberg (eds.), *Reason and Responsibility*, 10th ed. Belmont, CA: Wadsworth.
- Hume, David (1751/1998). *An Enquiry Concerning the Principles of Morals*, Tom L. Beauchamp (ed.), Oxford University Press.
- Hutcheson, Francis (1725/1991). *An Inquiry into the Original of our Ideas of Beauty and Virtue*. Reprinted in part in Raphael (1991), Vol. I, pp. 260–321. (First printed in 1725.)
- Jamieson, Dale (2002). "Sober and Wilson on Psychological Altruism." *Philosophy and Phenomenological Research* 65 (3): 702-710.
- May, Joshua (forthcoming). "Relational Desires and Empirical Evidence Against Psychological Egoism." *European Journal of Philosophy*.
- Maner, J. K., C. L. Luce, S. L. Neuberg, R. B. Cialdini, S. Brown, & B. J. Sagarin (2002). "The Effects of Perspective Taking on Motivations for Helping: Still No Evidence for Altruism." *Personality and Social Psychology Bulletin* 28(11), 1601-1610.

¹⁹ A version of this paper was presented at the 29th Spindel Conference at the University of Memphis, October 2010. I thank the participants for helpful feedback and stimulating discussion, especially Stephen Darwall, John Deigh, Remy Debes, Timothy Oakberg, Jesse Prinz, Tamar Schapiro, Walter Sinnott-Armstrong, and Michael Slote. Many thanks also to Dan Batson, Richard Holton, Ian Nance, and Aaron Zimmerman for helpful comments and discussion about this paper on other occasions.

- Monroe, Kristen R., M. C. Barton, & U. Klingemann (1990). "Altruism and the Theory of Rational Action: Rescuers of Jews in Nazi Europe." *Ethics* 101 (1): 103-22.
- Neuberg, S. L., R. B. Cialdini, S. L. Brown, C. Luce, B. J. Sagarin, & B. P. Lewis (1997). "Does Empathy Lead to Anything More than Superficial Helping? Comment on Batson et al." *Journal of Personality and Social Psychology* 73(3): 510-516.
- Oakberg, Timothy (manuscript). "A Critical Review of Batson's Project and Related Research on Altruism."
- Rosas, Alejandro (2002). "Psychological and Evolutionary Evidence for Altruism," *Biology and Philosophy* 17: 93-107.
- Slote, Michael A. (1964). "An Empirical Basis for Psychological Egoism." *Journal of Philosophy* 61(18): 530-537.
- Sober, Elliott & David Sloan Wilson (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.
- Stich, Stephen, John M. Doris, & Erica Roedder. (2010). "Altruism." Ch. 5 in *The Moral Psychology Handbook*, ed. by John M. Doris and the Moral Psychology Research Group, Oxford University Press, pp. 147-205.
- Williams, Bernard (1973). "Egoism and Altruism." Ch. 15 in *Problems of the Self: Philosophical Papers 1956-1972*, Cambridge University Press, pp. 250-265.