

## Empathy, Social Psychology, and Global Helping Traits

Christian Miller  
Wake Forest University  
[millerc@wfu.edu](mailto:millerc@wfu.edu)

Forthcoming in *Philosophical Studies*

The central virtue at issue in recent philosophical discussions of the empirical adequacy of virtue ethics has been the virtue of compassion. Opponents of virtue ethics such as Gilbert Harman and John Doris argue that experimental results from social psychology concerning helping behavior are best explained not by appealing to so-called ‘global’ character traits like compassion, but rather by appealing to external situational forces or, at best, to highly individualized ‘local’ character traits.<sup>1</sup>

In response, a number of philosophers have argued that virtue ethics can accommodate the empirical results in question, and have focused their attention in particular on explaining away the purported threat posed by the Milgram shock experiments, the Zimbardo prison experiments, and the Darley and Batson helping experiments.<sup>2</sup>

My own view is that neither side of this debate is looking in the right direction. For there is an impressive array of evidence from the social psychology literature which suggests that many people *do* possess one or more robust global character traits pertaining to helping others in need. But at the same time, such traits are a far cry from the traditional virtues like compassion. Thus at the risk of upsetting both sides of the

---

<sup>1</sup> See Doris 1998, 2002, Campbell 1999, and Harman 1999, 2000. The distinction between global versus local character traits will be developed in section one below.

<sup>2</sup> See, among others, DePaul 1999, Athanassoulis 2001, Sreenivasan 2002, Kamtekar 2004, and Sabini and Silver 2005 as well as my 2003. For the experiments, see Milgram 1963, Zimbardo 1973, and Darley and Batson 1973.

debate on the empirical adequacy of virtue ethics, I will argue that we should adopt an intermediate position between the situationist proposals of Harman and Doris and the leading responses being offered by virtue ethicists.

In what follows, I first summarize the central issues in the debate about the empirical adequacy of virtue ethics, and then in section two go on to develop the positive view of global helping traits. There I claim that one of the main ways of triggering such a trait is by empathizing with another person in need, and so section three characterizes empathy before some of the experimental evidence relating empathy and helping is presented in section four. The final two sections of the paper examine what bearing this view of empathy and global helping traits would have for virtue ethics if it turns out to be correct.

## 1. Background

In his recent book *Lack of Character*, Doris argues against the empirical adequacy of virtue ethics in particular and of any ethical theory more generally which ascribes a substantive role to global character traits. According to Doris, a *globalist* conception of character is one which accepts the following two theses in particular:

- (1) *Consistency*. Character and personality traits are reliably manifested in trait-relevant behavior across a diversity of trait-relevant eliciting conditions that may vary widely in their conduciveness to the manifestation of the trait in question.
- (2) *Stability*. Character and personality traits are reliably manifested in trait-relevant behaviors over iterated trials of similar trait-relevant eliciting conditions.<sup>3</sup>

---

<sup>3</sup> Doris 2002: 22. Doris also mentions a third globalist claim:

- (3) *Evaluative integration*. In a given character or personality the occurrence of a trait with a particular evaluative valence is probabilistically related to the occurrence of other traits with similar evaluative valences.

Thus according to this claim, a person who is honest, for example, would also be expected to have and manifest other character traits relevant to honesty, such as understanding, wisdom, and courage (Ibid).

Hence a *global character trait* is a character trait which exhibits both cross-situational consistency in a wide variety of trait-relevant circumstances, as well as iterated stability in repeated instances of the same kind of trait-relevant circumstances. Consider for instance how a trait such as honesty is understood according to such a globalist conception. Someone with this trait would be expected to behave honestly both in a wide variety of different honesty-relevant eliciting conditions (taking exams, testifying, talking to a spouse, etc.), as well as in repeated instances of similar conditions (i.e., many exams taken over multiple years). Such behavior would be no accident, as what would ground and hence explain why the person acts the way that he or she does in those circumstances is precisely a character trait of honesty.

Thus ascriptions of character traits to individual agents are supposed to play two central roles on a globalist framework – they are meant to explain consistent and stable manifestations of trait-relevant behavior, and they are supposed to accurately ground predictions of such behavior in the future. But, according to Doris, when we turn to empirical data in experimental social psychology, we find that situationism is the dominant research paradigm. Situationism rejects the first globalist thesis, and is neutral on the truth of the second.<sup>4</sup> Such a rejection stems from the kinds of experiments alluded to above, experiments which seem to show that behavior is highly influenced not by global traits but rather by a wide variety of situational influences. In particular, very few

---

However, evaluative integration is controversial even among virtue ethicists, and it is the first two conditions of consistency and stability which are crucial for Doris's critical discussion of the empirical adequacy of virtue ethics. Thus I leave this third condition to one side in what follows.

<sup>4</sup> More positively, situationism is characterized by Doris as a view which is committed to the follow three central claims:

- (a) "Behavioral variation across a population owes more to situational differences than disposition differences among persons . . .
- (b) Systematic observation problematizes the attribution of robust traits . . .
- (c) Personality is not often evaluatively integrated" (2002: 24-25).

people allegedly exhibit traits of character which are cross-situationally consistent; when someone is, say, honest in one situation, we find that he or she is often dishonest in a variety of other situations.

In the philosophical literature on situationism, Gilbert Harman seems at times to hold that the upshot of these empirical results in social psychology should be that there are no character traits whatsoever.<sup>5</sup> Doris, on the other hand, notes that situationism does not rule out the second thesis of globalism above, and argues that in fact there is evidence that people are remarkably consistent during temporal iterations of the same kind of situation. So he is willing to postulate the existence of very fine-grained 'local' character traits, traits which even if they exist nevertheless represent a significant departure from the global traits operative in traditional philosophical theorizing about character.<sup>6</sup>

So virtue ethics implies globalism, but globalism is incompatible with situationism, and situationism is empirically well-verified; therefore virtue ethics is empirically inadequate. As was noted above, support for this line of reasoning has been drawn almost entirely from certain studies on the helping behavior of subjects in social psychology experiments. But do these studies really rule out the existence of global traits pertaining to helping? It is not at all clear that they do, so long as we do not continue to think in terms of the virtue of compassion, but rather in terms of a different kind of character trait pertaining to helping others.

## **2. Global Helping Traits**

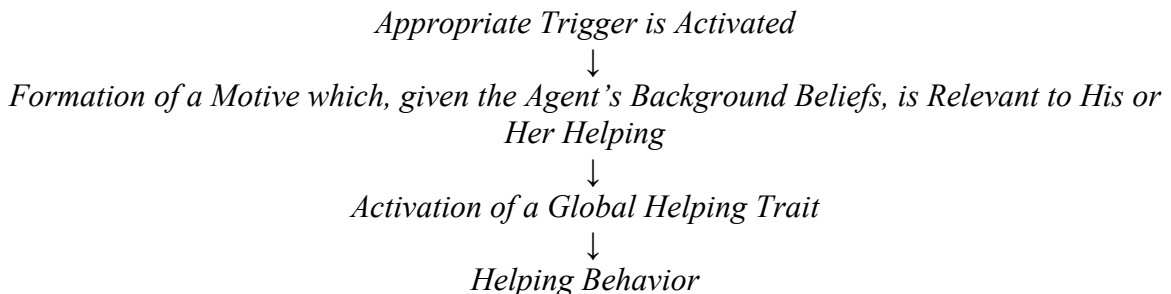
---

<sup>5</sup> See Harman 1999, 2000. The above has become a popular way of understanding Harman's view, but to be fair at other times he only seems to be rejecting the existence of what he calls 'broad-based' dispositions, i.e., traits of character which meet Doris' first criterion for being global. Fortunately for our purposes nothing hangs on which interpretation proves to be correct.

<sup>6</sup> Doris 2002: 23, 25, 64.

On my reading of the social psychology literature, the empirical results are compatible with the widespread existence of one or more global character traits associated with helping. On the other hand, we will see in section six of the paper that such a result is of little consolation to virtue ethicists since even if they do exist, such traits are rather different from how compassion has traditionally been construed. So in order to prevent them from being conflated with compassion, let us call these traits ‘global helping traits’ or GHTs.

GHTs are dispositions to help others who are thought to be in need. In addition, they are highly sensitive to a number of different psychological inputs such that a GHT will be ‘triggered’ and will lead the agent to try to help, other things being equal, provided that one or more of these inputs is sent to the GHT at a suitable strength to pass above its minimal activation threshold. Diagrammatically, the picture is as follows:



where the arrows are intended to symbolize causal influence. Thus as we will see in a moment, one such trigger might be feelings of guilt, which can cause the formation of a motive to relieve the guilt. Given the agent’s background beliefs that helping is a means of relieving guilt and that the agent is in a position to help certain people, this motive could activate the relevant GHT which in turn leads to the agent trying to do so. But note that this need not be the only kind of motive at work in our model – depending on the kind of trigger in question, others might be, for instance, a motive to reduce another’s

distress, or to maintain one's good mood, or to relieve a bad mood. In each case it is easy to see how, depending on the agent's beliefs, such motives could lead to the activation of a helping mechanism.<sup>7</sup>

Admittedly, given the very modest levels of helping behavior actually exhibited by control subjects in social psychology experiments, we have good reason for thinking that in ordinary circumstances GHTs do not get triggered and so typically do not play a robust causal role in many peoples' psychological lives. But once activated in one of the relevant ways, global helping traits can directly lead a person to help in a wide variety of circumstances. Similarly, so long as it remains activated or is activated again, a GHT can lead to repeated attempts to help in similar circumstances. So such a trait bears the two hallmarks of a global character trait – it can give rise to both cross-situationally consistent and iteratively stable behavior.<sup>8</sup>

---

<sup>7</sup> It is important to note that the diagram above is only intended to apply to *paradigm* cases of helping which stem from a GHT. GHTs can lead to helping behavior even without one of the relevant triggers being activated, say as a result of a standard case of practical reasoning which concludes with a motive that is relevant to helping a particular person. Even a wayward causal chain in the agent's psychology or direct stimulus from the environment might conceivably form a motive which is taken by the agent to be relevant to his or her helping and which in turn activates a GHT. Similarly, a particular agent might only be psychologically capable of having certain triggers be activated and not others. An agent who does not experience guilt, for example, could still have a GHT and have it be triggered in other ways.

Thanks to Nancy Snow for pressing me to address these issues.

<sup>8</sup> It is important to note that the proposal is *not* that GHTs can give rise to helping behavior in a given situation only when appropriately triggered *in that situation*. Rather, a GHT can be triggered in one situation, and still give rise to helping behavior in a number of subsequent situations (whether of the same or disparate type) so long as it remains activated.

Thus strictly speaking, GHTs can only exhibit cross-situational consistency *once initially triggered*. As a result, it may not be clear whether this property of a GHT is compatible with the letter of Doris' consistency requirement on global character traits (and corresponding remarks apply in what follows to the stability requirement as well), since it might be claimed that GHTs would *not* be reliably manifested in trait-relevant behavior across a diversity of trait-relevant eliciting conditions, but would only be so manifested once triggered. My own suspicion is that GHTs are compatible with Doris' formulation of the requirement, since even virtues as traditionally understood have to be triggered by certain relevant inputs such as a belief that someone is in need (certain cases of compassion), or that someone is in danger (certain cases of courage), or that something is unhealthy (certain cases of temperance). And naturally Doris wants all traditional virtues to count as global character traits.

However, even if GHTs are not strictly compatible with the wording of Doris' requirement, they are clearly consistent with the *spirit* of a globalist conception. For such GHTs are robust psychological structures which, when appropriately triggered, would form a significant part of the causal explanation for

What are these triggers to which GHTs are allegedly sensitive? The past forty years of research in social psychology have shown that helping behavior is remarkably sensitive to the following psychological factors (among others):

- Guilt.<sup>9</sup>
- Embarrassment.<sup>10</sup>
- Moderate Good Moods.<sup>11</sup>
- Moderate Bad Moods.<sup>12</sup>
- Empathy.<sup>13</sup>

For instance, literally hundreds of studies have shown that subjects who are put into a moderately good mood are much more likely to help. Thus in Weyant's well-known 1978 experiment some subjects had their affect levels raised by being made to believe that they had performed well on a fake anagram test. After learning the results of the test, they were presented with an opportunity to donate their time to charity work. Of the 252 subjects, random assignments were made as to which of them would be presented with one of the following opportunities:

- American Cancer Society (high benefits) and Door-to-Door Work (high costs)
- American Cancer Society (high benefits) and Desk Work (low costs)
- Little League Baseball (low benefits) and Door-to-Door Work (high costs)
- Little League Baseball (low benefits) and Desk Work (low costs)

The percentage of subjects who volunteered came out as follows:<sup>14</sup>

	Positive Affect	Controls
High Benefits / High Costs	57%	33%
High Benefits / Low Costs	62%	33%
Low Benefits / High Costs	52%	29%

the performance of helping behavior in a wide variety of circumstances, as well as ground accurate predictions of such future actions. Furthermore, even if such traits are not strictly speaking 'global,' their existence is clearly incompatible with Doris' proposal to only accept local traits, as well as with Harman's rejection of all character traits whatsoever (or even just those which are 'broad-based' – see footnote five).

I am grateful to an anonymous referee for pressing me to clarify these issues.

<sup>9</sup> See Regan 1971.

<sup>10</sup> See Apsler 1975.

<sup>11</sup> See Isen 1987, Carlson, Charlin, and Miller 1988, and Schaller and Cialdini 1990.

<sup>12</sup> See Cialdini and Kenrick 1976, Manucia, Baumann and Cialdini 1984, Cialdini et al. 1987, and Cialdini and Fultz 1990.

<sup>13</sup> See Batson 1987, 1991, 2002, and Batson et al. 2003.

<sup>14</sup> Weyant 1978: 1173.

Low Benefits / Low Costs	62%	33%
--------------------------	-----	-----

When it comes to *bad* moods, subjects will also help more than control subjects, but only when they perceive the benefits of helping to outweigh the purported costs. We can see this as well using Weyant’s study when different subjects were told they had performed badly on the test. This time the percentages of subjects who volunteered their time were:<sup>15</sup>

	Negative Affect	Controls
High Benefits / High Costs	29%	33%
High Benefits / Low Costs	71%	33%
Low Benefits / High Costs	5%	29%
Low Benefits / Low Costs	33%	33%

Other studies have shown that much higher levels of helping behavior were also exhibited by subjects who experienced elevated levels of guilt, embarrassment, or empathy, the last of which will be examined in greater detail in section four.

Assuming that most people possess global helping traits which are sensitive to the above triggers, we should be able to predict patterns of cross-situational consistency and iterated stability. When it comes to mood states, for example, we should expect that, other things being equal, many people with elevated levels of positive affect will be such that they help others in situations ranging from picking up dropped papers to volunteering time for charity organizations. On the other hand, we can expect that, again other things being equal, many people without elevated levels of affect will be such that they do not exhibit these and other forms of helping behavior. And in study after study, these are the patterns we do in fact find.

More precisely, we can formulate conditionals which give empirical predictions for helping behavior in order to test whether subjects have such global helping traits.

---

<sup>15</sup> Ibid.

Since as Doris himself notes, “sporadic failures of trait-relevant behavior probably shouldn’t be taken to disconfirm attributions,”<sup>16</sup> we can build probabilistic qualifiers into the consequents of the conditionals. Thus for positive affect, we would get conditionals like the following:

- (a) If an adult possesses a GHT and is experiencing intermediate levels of increased positive affect, that person will probably engage in helping-relevant behavior in moderate helping-relevant eliciting conditions.

The ‘moderate’ qualifier in the consequent is intended to exclude what are taken by the agent to be extremely demanding acts of assistance, which we can predict are not likely to be performed very frequently and which are not at issue in discussions of mood effect studies.<sup>17</sup>

Turning to negative affect, the experimental data indicates that subjects will help but only as a means of improving their moods, and only provided that there are no more beneficial courses of action to take.<sup>18</sup> Thus we can predict that:

- (b) If an adult possesses a GHT and:
  - (i) Is experiencing intermediate levels of increased negative affect.
  - (ii) Takes the *benefits* for himself of helping to outweigh the perceived costs to himself.
  - (iii) Does not take there to be any more effective means available for relieving the negative affect.that person *will* probably engage in helping-relevant behavior in moderate helping-relevant eliciting conditions.

Similarly for a trigger like empathy:

- (c) If an adult possesses a GHT and is experiencing intermediate levels of increased empathy, that person will probably engage in helping-relevant behavior in moderate helping-relevant eliciting conditions.

However, if no inputs are present to trigger a GHT, then:

- (d) If an adult possesses a GHT which has not been triggered, that person will probably not engage in helping-relevant behavior in moderate helping-relevant eliciting conditions.

---

<sup>16</sup> Doris 2002: 19.

<sup>17</sup> As Doris agrees (2002: 49).

<sup>18</sup> For references, see footnote twelve.

Naturally it is assumed that various other relevant considerations are being held equal, i.e., that the person is not also experiencing depression or an intense emotion like anger or fear.

Thus far we have been focusing on the way in which the activity of a GHT might be enhanced when triggered in one or more of the ways listed above. While not the main concern of this paper, it is worth briefly noting that such traits are also sensitive to a wide variety of psychological influences which serve as *inhibitors* rather than enhancers of their functioning properly. For example, subjects in a bad mood who take the costs to themselves of helping to significantly outweigh the benefits of doing so help much less than control subjects, as can be seen with the 5% versus 29% helping rates in Weyant's study. Similarly, it is a well-documented phenomenon that the presence of other people serves as a powerful inhibitor to helping behavior.<sup>19</sup> For example, Latané and Nida examined 37 studies comparing subjects by themselves in emergencies versus subjects in groups and found that 50% of subjects helped in the first kind of situation whereas only 22% did so in groups.<sup>20</sup> And it is a striking feature of the bystander effect literature that the results are normally presented in terms of inhibiting helping dispositions rather than as evidence for the nonexistence of helping traits.<sup>21</sup> With such results in mind, a GHT would then seem to be remarkably sensitive to certain bad moods and to the presence of bystanders as well as to a host of other inhibitors such as anger and depression.<sup>22</sup>

---

<sup>19</sup> Latané and Darley have outlined three psychological processes – audience inhibition, social influence, and diffusion of responsibility – in an attempt to fully explain the way helping is inhibited in social contexts. See their 1970 and the very helpful overview in Latané and Nida 1981.

<sup>20</sup> Latané and Nida 1981: 321.

<sup>21</sup> I have been helped here by Nancy Snow.

<sup>22</sup> For anger, see Cialdini, Baumann, and Kenrick 1981. In a more extensive treatment of GHTs, their alleged existence would have to be assessed in light of the three experiments mentioned at the start of this paper – the Milgram shock experiments, the Zimbardo prison experiments, and the Darley and Batson

What I am claiming about GHTs should not, however, be overstated. Strictly speaking, all I can legitimately assert is that the social psychology literature is *compatible* with the widespread possession of GHTs by normal adult human beings. It does not, however, offer direct evidential *support* for the possession of such traits. This is because in order to be able to reliably test the extent to which people possess GHTs, if they do so at all, we would need to perform longitudinal experiments which track subjects over time as they find themselves in both similar and different help-eliciting circumstances. Only in such a manner can we see whether their behavior is consistent and stable. Unfortunately, because of the financial and logistical difficulties associated with carrying out these experiments, no such empirical evidence is available at the present time, or at least not as far as I am aware. Instead, the experiments that we have in this area only look at how subjects act at a particular time and place.

Two things might nonetheless be said in order to strengthen the claim being made about GHTs. First, Doris himself laments the unavailability of the longitudinal data which would be most relevant for evaluating the existence of global character traits like compassion or honesty.<sup>23</sup> Thus without this data, one might think that all Doris is also entitled to claim is that, even if his controversial interpretation of the experimental literature is correct, the data that we have now is only compatible with the non-existence of global character traits. So if I too am merely making a compatibility claim concerning

---

seminary helping experiments. Briefly, in each case we could cite a plausible factor – obedience and conformity pressures, bystander effects, and distraction and hurry influences respectively – which could have inhibited the activation and/or proper functioning of the subjects’ GHTs. Clearly the role of potential inhibitors would be featured much more prominently in such a discussion, although it is best left for another paper. Similarly, other sources of alleged support for the situationist rejection of global traits associated with helping, such as real world destructive behavior in the form of the Holocaust or the Gulag, would also need to be examined in detail. So even if this paper is correct in arguing that the existence of GHTs is compatible with large segments of the psychology data on helping, a much longer discussion is needed in order to argue for a compatibility claim across the board.

<sup>23</sup> Doris 2002: 38.

global character traits and the social psychology literature, then such a claim is as strong as what anyone in this area is entitled to assert.

But secondly, with over forty years of work and the results of hundreds of helping experiments in place, many social psychologists today make predictions about the outcomes of their future experiments which exhibit a commitment to conditionals that would have to be grounded in relatively stable helping structures. And this is presumably because even despite the absence of longitudinal data, they take the existing experiments to constitute good inductive evidence for consistent patterns of behavior which they believe will also be exhibited in novel experimental conditions. In other words, social psychologists seem justified in predicting that (other things being equal) the majority of subjects would help in certain situations which are different from those that have already been tested, provided the right psychological triggers for helping are also in place. And it routinely turns out that such helping behavior does result, thereby confirming their hypotheses. Given the variations in the situations and yet the same consistent patterns of helping behavior, it is only natural to assume that one or more stable helping structures is operative behind the scenes, and to base future predictions accordingly. So perhaps something stronger than a mere compatibility claim might be warranted after all.

Regardless of whether this last point is defensible or not, I mainly want to suggest that we have good reason to *not* take the social psychology literature to rule out the existence of global traits pertaining to helping. The above claims have, however, been rather abstract, and so we shall turn in the remainder of the paper to a more detailed and empirically informed discussion of the impact that a specific input, namely empathetic feelings, can have on GHTs and subsequent helping behavior.

### 3. Characterizing Empathy

First we will need to have some idea of what empathy involves. Given limitations of space I will not attempt anything approaching an analysis of the concept here. Nor will I be concerned with trying to delineate the different modes of empathetic arousal experienced by infants, young children, and adults that have recently been distinguished in the developmental literature on empathy.<sup>24</sup> Rather, my concern will merely be with drawing attention to some of the central features of empathy in paradigm cases involving normal adult humans. Furthermore, these will be cases involving empathetic *feelings*; one might be able to empathize with how a person came to form certain beliefs or be misled by the evidence, but this kind of empathy will not be our concern here.<sup>25</sup>

A paradigm case of empathy for the feelings of another might go as follows. John is good friends with Jennifer. Jennifer suddenly loses her parents, and is experiencing tremendous distress. John tries to imagine what she must be feeling, and as a result, comes to form similar feelings in his own mind. In this way, John has come to empathize with what Jennifer is going through.<sup>26</sup>

The first thing that is important to note about this case is what John is supposed to be imagining when he is thinking about Jennifer's distress. Here we can distinguish at least two proposals:

- (i) John tries to imagine what *Jennifer* perceives in the situation and what she feels as a result.
- (ii) John tries to imagine what *he* would perceive in the situation if he were in Jennifer's position and how he would feel as a result.

---

<sup>24</sup> For overviews, see the papers in Eisenberg and Strayer 1987a and Baron-Cohen 1993.

<sup>25</sup> In the remainder of this section I have been helped most by Snow 2000. For related discussion, see also the papers in Eisenberg and Strayer 1987a as well as Sober and Wilson 1998 and Darwall 1998.

<sup>26</sup> For a similar example, see Snow 2000: 66.

These are clearly two different acts of imagining, and can give rise to noticeably different feelings. For instance, Batson, Early, and Salvarani (1997) have shown that the second act of imagining leads to feelings of personal distress in the subject which are not typically found in those who actively imagine in the first way. And it seems clear on intuitive grounds that it is the first kind of imagining which is conceptually tied to empathy.<sup>27</sup>

The second thing to note about paradigm cases of empathy is that someone like John does not have to feel exactly the same thing that Jennifer does in order to be said to have empathized with her situation; rather he needs to only have felt an emotion which is similar in kind to hers. Thus John does not have to currently experience quite the degree of distress as Jennifer does, and may only be feeling sadness in comparison to her deep depression.<sup>28</sup>

Third, although the point may be obvious, it is worth emphasizing that John forms a feeling empathetically in virtue of *believing* that Jennifer is experiencing a similar feeling. Such a belief is perhaps not necessary in all cases of empathy – young children can have their empathetic feelings caused by emotional contagion in which they are said to ‘catch’ a feeling of a parent directly from the parent’s body language, facial expressions, or tone of voice.<sup>29</sup> However, focusing as we are on paradigm cases involving imaginative contemplation of another’s mental life, an agent won’t be able to empathize

---

<sup>27</sup> As has been repeatedly noted in the empathy literature. See, e.g., Batson 1987: 93, Wispé 1986, Batson, Early, and Salvarani 1997, Darwall 1998, and Snow 2000.

<sup>28</sup> For similar claims, see Sober and Wilson 1998: 233 and Snow 2000: 69.

<sup>29</sup> For more, see the discussions of infant and childhood emotional development in Eisenberg and Strayer 1987a, Baron-Cohen 1993, and Darwall 1998: 264-266.

with the feelings that another is supposed to have without first believing that she has them to begin with.<sup>30</sup>

Our final point is that it is important to not conflate empathy with sympathy. Sympathy is an emotion which involves some form of care or concern for another person. The other person is the object of this state, and so the attitude is third-personal rather than first-personal. As we have seen, empathy is rather different – one adopts the first person perspective of the other person, and thinks about the world *with* her, rather than being directly concerned *about* her. Thus in our example, John feels sadness *with* Jennifer, which may or may not give rise to a separate sympathetic emotion *for* her. By empathizing with Jennifer, John becomes focused with her on the death of her parents. By sympathizing with Jennifer, John becomes focused on something else, namely Jennifer herself.<sup>31</sup>

We can summarize these points about empathy as follows:

If John empathizes with Jennifer's feeling, F, then:

- (i) He tries to imagine what Jennifer feels in the situation, rather than what feelings he would have if he were in the situation.
- (ii) He only needs to experience something similar to F, rather than F itself, to empathize with Jennifer's feeling.
- (iii) He believes that Jennifer is feeling F (or something similar to F).
- (iv) His empathizing with Jennifer's feeling F is conceptually distinct from his sympathizing with her.

Clearly much more could be said in defense of these conditions and about empathy in general, but hopefully they serve their primary purpose, which is to clarify the phenomenon at issue in the remainder of this paper. And I am not going to fight about

---

<sup>30</sup> See Sober and Wilson 1998: 234 and Snow 2000: 68. Furthermore, whatever John does feel, he must also feel it *because* Jennifer has a similar feeling; in other words, Jennifer's having such a feeling must have been what ultimately gave rise to John's affective experience. Otherwise, he could have just formed that feeling by chance, or because of a deviant causal process in his psychology which led him to form a belief about Jennifer's feelings. Such etiologies would preclude the relevant feeling from counting as empathetic. For additional discussion, see Snow 2000: 65-67.

<sup>31</sup> For similar views, see Wispé 1986: 318, Eisenberg and Strayer 1987b: 5-6, Darwall 1998, Snow 2000: 66, and Slote 2004: 299.

labels; if the above does not sound like ‘empathy’ then we can give it a different name and still see whether subjects who are in this state of mind are more disposed to help those whom they think are in need.

#### **4. Empathy and Helping Behavior**

Before we turn to the literature on empathy and helping behavior, a natural worry that should be addressed from the start is whether the experiments that have been performed by social psychologists are appropriately sensitive to the distinctions mentioned in the previous section and so are careful enough to test for empathy rather than, say, sympathy or other related emotional states. At least in the case of the leading psychologists working in this area, such as Daniel Batson and his colleagues, it turns out that the experiments have been appropriately sensitive. In order to induce empathy in the undergraduate volunteers who served as subjects, Batson had them read or listen to accounts of people experiencing some kind of hardship. For instance, one (fictional) case involved a radio broadcast about Katie Banks, a university student whose parents and a sister had been killed in a car crash and who was left to care for her younger brother and sister. The empathy manipulation was achieved beforehand by having subjects read a passage such as the following:

While you are listening to this broadcast, try to *imagine how the person being interviewed feels about what has happened and how it has affected his or her life*. Try not to concern yourself with attending to all the information presented. Just concentrate on trying to imagine how the person interviewed in the broadcast feels.<sup>32</sup>

Note that the aim of these instructions was to directly stimulate empathy, not sympathy, and that subjects who felt something similar to what Katie felt did so because of beliefs they had about her feelings. Furthermore, Batson’s experiments were sensitive to whether

---

<sup>32</sup> Batson, Early, and Salvarani 1997: 753, emphasis theirs.

the subjects should imagine their own feelings or those had by Katie in the situation. For he has explicitly distinguished the above set of instructions from these:

While you are listening to this broadcast, try to *imagine how you yourself would feel if you were experiencing what has happened to the person being interviewed and how this experience would affect your life*. Try not to concern yourself with attending to all the information presented. Just concentrate on trying to imagine how you yourself would feel.<sup>33</sup>

Finally, both of these instructions have been distinguished from those given to control subjects whose actions have served as the baseline level of exhibited helping behavior:

While you are listening to this broadcast, try to *be as objective as possible about what has happened to the person interviewed and how it has affected his or her life*. To remain objective, do not let yourself get caught up in imagining what this person has been through and how he or she feels as a result. Just try to remain objective and detached.<sup>34</sup>

So it seems as if the highly influential experiments in this area by Batson and his colleagues do involve instructions which are sensitive enough to some of the central features of empathy to be able to generate relevant results.

What have these experiments shown thus far? More than fifty different experiments involving many different feelings of distress, helping opportunities, and helping tasks have shown that subjects who are induced to feel empathy exhibit significantly more helping behavior than control subjects towards those thought to be in need. Here is not the place to examine all of these studies,<sup>35</sup> but as a representative example we can consider the following from Toi and Batson (1982). Half of the volunteers from an introduction to psychology course were asked to listen to a broadcast and be as objective as possible, whereas the other half were told to imagine the perspective of the person being interviewed. The tape they each heard next contained a

---

<sup>33</sup> Ibid., emphasis theirs.

<sup>34</sup> Ibid., emphasis theirs.

<sup>35</sup> For reviews, see Batson 1987, 1991, 2002, and Batson et al. 2003. For other related studies, see the references in Eisenberg and Miller 1987: 300-301 and Dovidio and Penner 2001.

(fictional) interview with Carol Marcy, a freshman in the class who had had both of her legs broken in an auto accident and was worried about being able to still pass the course. After listening to the interview and filling out a questionnaire, subjects received an envelope with letters from both the professor of the course and from Carol asking for help in going over the missed lecture notes. Furthermore, Carol indicated that she would not be coming back to school but would be studying at home until next semester. Thus Toi and Batson reasoned that subjects who did not offer to help would not have to actually see Carol for the rest of the semester and so would not feel (as) guilty. The dependent measure was whether the subjects filled out a slip agreeing to help Carol. Here were the percentages who volunteered:<sup>36</sup>

Controls	33%
Empathized	71%

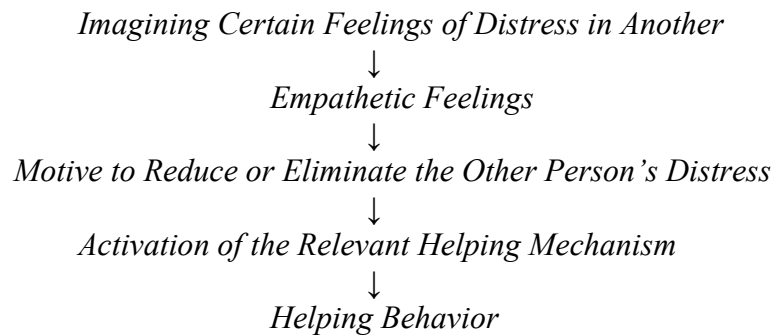
The only relevant difference in the experimental setup was a difference of two sentences in the instructions the two groups were given before listening to the broadcast, and yet the different perspectives that were thereby generated led to dramatic differences in actual helping behavior. As indicated above, similar patterns have arisen in dozens of other experiments conducted by Batson and replicated by other social psychologists.

Given these results, we can construct a crude but accurate picture of what seems to be going on in experiments such as the above. Subjects are instructed to adopt an empathetic stance at some point in the near future, and when the time comes they actually do empathize with people such as Carol and form empathetic feelings because of what they believe the other person is feeling in her situation. These empathetic feelings in turn generate a motive to reduce or eliminate the pain, suffering, distress, or other relevant

---

<sup>36</sup> Toi and Batson 1982: 288.

difficulty that the person is experiencing.<sup>37</sup> Helping is naturally a way of trying to eliminate the problem, and so relevant helping mechanisms are activated and the likelihood that the subject will help is increased.<sup>38</sup> Thus diagrammatically we have the following:



where again the arrows are intended to symbolize causal influence. In the next section we will say much more about the motive to eliminate the other person's distress. For now the main point to take away from this model is that many subjects in the relevant psychology experiments are such that empathy for a person deemed to be in need significantly augments helping behavior.

Assuming that the experimental results can be trusted, how do they relate to the larger concerns of this paper? What we are calling the relevant helping mechanisms in the above diagram are just the global helping traits or GHTs that were discussed at length in section two.<sup>39</sup> That there are such traits, and that they can be triggered by, among other

---

<sup>37</sup> Batson 1987: 95.

<sup>38</sup> For similar models, see Coke, Batson, and McDavis 1978, Darwall 1998, and Dovidio and Penner 2001: 184.

<sup>39</sup> As noted in footnote seven, this diagram is only intended to illustrate a paradigm instance of the way in which empathetic feelings might lead to the activation a GHT. GHTs can be activated in a number of other ways, and can exist even in subjects who are incapable of empathetic feelings. Similarly, as we noted at the start of section three, there are several different modes of empathetic arousal, and the focus of this paper has been only on the kind of empathetic feelings typically found in normal adult humans. So nothing I say should be taken to preclude the possibility that believing that another is in distress (as opposed to imagining the feelings of that person) might also foster motives which lead to the activation of a GHT. Even the kind of emotional contagion in which people simply pick up on the distress of others might produce a motive

things, motives generated by empathetic emotional experiences, are claims that should now seem to be compatible with the sizeable literature on the role that empathy has to play in augmenting helping. Of course, we need to keep in mind that because this literature does not include any longitudinal studies, it is difficult to make any stronger claims beyond mere compatibility with the experimental results.<sup>40</sup> At the same time, it is admittedly hard to see what could explain the strong correlation between empathy and helping if some kind of global helping trait did not exist. Situational forces alone do not seem as if they could do the job – the experiments were conducted using many different subjects, different people in need, and different opportunities to help, and yet exhibited the same pattern throughout. Similarly, it would be remarkable if these undergraduate subjects had happened to have already cultivated local helping traits for the different situations they were placed in which disposed them to help such that over a span of thirty years the same correlations between empathy and helping behavior always emerged.<sup>41</sup>

Rather it is far more natural to think that some kind of global helping structure was in place, one which would show significant cross-situational consistency and iterated stability provided it was repeatedly triggered in the right way.<sup>42</sup> Such a global structure would be the basis for an explanation as to why – despite a wide range of subjects, people in need, and helping tasks – the same consistent trend of augmented helping behavior was found in empathy induced subjects but not in control subjects. And such a GHT would

---

which in turn activates a GHT. Indeed I suspect that experimental results could be produced for such kinds of empathy which show trends that are similar to those examined in this section. Again I am grateful to Nancy Snow for pointing out the need to clarify these issues.

<sup>40</sup> However, see Oliner and Oliner 1988, who found that forty years after World War II, people who had saved Jews from the Nazis were still more helpful than those who had not. For other suggestive longitudinal studies, see Dovidio and Penner 2001: 182-184.

<sup>41</sup> This point will be developed at greater length in the next section.

<sup>42</sup> As two psychologists who work on empathy note, “if an empathic response engenders sympathy for the other, one would expect an association between empathy and prosocial behavior across a wide variety of situations” (Eisenberg and Miller 1987: 296).

serve as the ground for making predictions of such behavior in the future when empathy induced subjects are placed in novel environments or are given new opportunities to help. If the experimental results continue to be in line with what GHTs would lead one to expect, then absent actual longitudinal studies we would have the closest thing to empirical support for these traits that we can manage. And given the dozens of studies like this that have already been conducted, there is every reason to be confident that future results *will* continue to show similar patterns of helping behavior.

## **5. The Good News for Virtue Ethics**

With this brief sketch in mind of recent work in social psychology on empathy and helping, what should we conclude about the viability of virtue ethics and the success of Harman and Doris's general line of criticism? Overall the results do not seem to obviously favor either side, and instead an intermediate position will emerge between the two opposing views. In this section, we shall see why the data suggests that two fundamental pieces of the virtue ethicist's conception of compassion are not empirically threatened – the claim that compassion must be a global character trait, and the claim that compassionate motives must be altruistic rather than self-interested. However, in the next section we shall see that these claims are not enough to put to rest all of the concerns that Harman and Doris have raised.

*Global Character Traits.* As has been noted several times already, Harman and Doris overstate the conclusions one can draw from the social psychology literature; the experimental evidence does not rule out the widespread possession of a global character trait like the GHTs from section two. The mere fact that virtue ethicists are committed to

compassion's being a global as opposed to a local trait is, by itself, no source of embarrassment.

In addition we do not find any support for local character traits in this part of the helping literature. Recall that as an alternative to globalism about character, Doris proposed a view according to which there are a number of fine-grained local traits which have iterative stability but not cross-situational consistency. Such traits are situation particular, and while Doris does not give any rigorous conditions for individuating situations, he does say that he intends them to be differentiated "in terms of environmental features characterizable independently of individual psychological particularities."<sup>43</sup>

If our characters merely consisted of a collection of various local traits, we would expect helping behavior to be highly fragmented – many people experiencing empathy might volunteer to tutor Carol in psychology but not to drive her to school, while a significant number of those experiencing no raised levels of affect might, for instance, make change for a dollar but not donate blood. Yet this is not in fact the way that the behavioral patterns have turned out – the empirical data suggests that a significant number of people with raised levels of empathetic emotion would help in a wide variety of circumstances, while many of those without such raised levels would not. So there seems to be more structure at work than a fragmentation model of character would lead us to believe.<sup>44</sup>

---

<sup>43</sup> Doris 2002: 76.

<sup>44</sup> Doris might respond that given the similarities between the situations most of us end up confronting, situational forces have habituated us into having roughly the same set of local traits associated with helping. Admittedly such a response would account for the results, but it also seems rather difficult to believe. For is it really plausible to think that many of us have, through a process of gradual habituation, acquired separate traits for picking up dropped papers, making change, donating blood, and volunteering for charity work? After all, it is not even clear that many people have been exposed to even a few, much less a significant

Admittedly, these claims must be made with a great deal of hesitancy without detailed longitudinal studies. So at best perhaps what we should conclude is that the studies on empathy and helping do not offer any support for the existence of local helping traits, but at the same time they do not definitively rule them out either.

*Altruistic Motives.* According to traditional forms of Aristotelian virtue ethics, virtuous agents will not only perform right actions but also do them for the right motivating reasons. Thus someone who helps another might be doing what he ought to do, but if he does it either solely or even in large part because of considerations such as social recognition or monetary reward, he would not be exhibiting the virtue of compassion.

Returning to the model of empathy and helping behavior we saw at the end of the previous section, we said that imagining the feelings of another in distress can lead to the formation of a motive to relieve that person's distress. But so far nothing in this paper has indicated whether that motive is egoistic or altruistic. Fortunately there has been a wealth of experimental evidence provided by Batson and his colleagues for what he calls the empathy-altruism hypothesis, or the claim that "empathy evokes motivation directed toward the ultimate goal of reducing the needy person's suffering; the more empathy felt for a person in need, the more altruistic motivation to have that need reduced."<sup>45</sup> In order to defend this claim about altruistic motivation, Batson has had to examine a number of possible egoistic explanations for the motivational impact that empathy has on helping behavior, and has devised separate experimental approaches for testing most of them. In every case, no empirical support has been found for any of the egoistic explanations.

---

number of repeated instances of these situation types so that they could have inculcated the relevant local character trait through habituation. Thus it seems that we would be left with a mystery as to how we have come to acquire such discrete and fine-grained helping traits in the first place.

<sup>45</sup> Batson 2002: 92.

Clearly an evaluation of Batson's treatment of each of these candidate explanations is beyond the scope of this paper, but we can get an initial appreciation of his results by combining the egoistic explanations into three categories and seeing what a representative experiment looks like for testing each of them.<sup>46</sup>

(i) Aversive Arousal Reduction. According to this explanatory strategy, feeling empathy is thought to be unpleasant or distressful for the person experiencing it, which in turn generates motivation to end such feelings. One such means to do so is to help the person in need and so the subject is motivated to help, but only as a way of making himself feel better.<sup>47</sup> If this explanation is correct, then the resulting helping behavior clearly cannot be said to have arisen from a compassionate character trait.

One way to experimentally test this proposal is to provide subjects with an opportunity to escape being around the person in need. If they are primarily motivated by seeking ways of reducing their aversive arousal, then they will take the opportunity to escape without helping. But this is not what has happened in the experiments. Recall, for instance, Toi and Batson's study in which subjects were told to either objectively or imaginatively listen to a broadcast which turned out to be about Carol's auto accident. The experiment had an additional wrinkle which was not mentioned earlier, namely that only some subjects were told that Carol would not be coming back to their psychology

---

<sup>46</sup> The three categories used follow Batson 2002: 94. See also Batson 1987: 84, 105 and Batson et al. 2003: 281-284. For reviews of the experimental support for the empathy-altruism hypothesis, see Batson 1987, 1991, 2002, and Batson et al. 2003. One explanation that does not fit nicely into these categories is the oneness hypothesis advanced by Cialdini et al. 1997, according to which by empathizing with another person, one incorporates one's self in the other, and the division between self and other breaks down. This in turn undermines the distinction between altruism and egoism. For strong empirical disconfirmation of this proposal, see Batson et al. 1997. Finally, Stich et al. 2007 examine Batson's work in detail and suggest that additional experimental results are needed before we can completely rule out certain egoistic explanations.

<sup>47</sup> For discussion, see Batson et al. 1981, Toi and Batson 1982, Batson et al. 1983, Cialdini et al. 1987, Batson et al. 1989, Batson 2002: 94-95, and Batson et al. 2003: 284.

class. Others were given a more difficult to escape scenario in which they were told that Carol would be attending all the remaining class, and given that she would be in a wheelchair, she would be hard to avoid. Here is how the proportion of subjects helping in each condition turned out:<sup>48</sup>

<i>Ease of Escape</i>	<i>Objective Instructions</i>	<i>Empathy-Inducing Instructions</i>
Easy	33%	71%
Difficult	76%	81%

Thus subjects feeling empathy for Carol did not appear to be significantly motivated by finding ways to reduce their feelings of distress, unlike subjects who took a more detached perspective. Similar results have been produced in variants of this experimental setup.<sup>49</sup>

(ii) Empathy-Specific Punishment. Another family of egoistic explanations maintains that people who feel empathy for those in need help primarily in order to avoid one or more punishments for not helping. Such punishments can range from social or religious disapproval to forms of self-censure involving guilt or shame.<sup>50</sup> Thus again people would not be ultimately compassionate towards those in need in such cases since they would be concerned mainly about themselves and what they can do to avoid the relevant form of punishment.

One way to test at least some explanations which fall under this category is to compare the change in mood experienced by subjects when told that their efforts to help someone in need failed. If empathy-specific punishment models are accurate, then there

---

<sup>48</sup> Toi and Batson 1982: 288.

<sup>49</sup> See, e.g., Batson et al. 1981, Batson et al. 1983, and Batson et al. 1989.

<sup>50</sup> For discussion, see Batson et al. 1988, Batson 2002: 95-96, and Batson et al. 2003: 282-284.

should be a noticeable difference in the mood change between subjects who were told that their failure to help was unjustified, and those who were told that it was justified. In particular, if subjects were told that their failure was justified, then we would expect them to be greatly relieved, whereas if it was unjustified, then they should be highly distressed. If the empathy-altruism hypothesis is correct, however, there should be little to no mood change between justified versus unjustified failures as the person is concerned only with helping the other in need, and the need still has not been addressed.<sup>51</sup>

Batson and Weeks (1996) carried out an experiment with just this design. The (fictional) person in need was Julie, who would be receiving a series of mild but uncomfortable shocks. After half the subjects were given either the usual set of objective or empathy-inducing instructions regarding an upcoming communication, they listened to a tape in which Julie expresses her anxiety about having to receive the shocks. An emotion survey was then filled out, followed by a task which, if performed successfully, could have saved Julie from the shocks. Some subjects were told afterwards that they failed the task and it had been “Moderately Easy,” whereas others were told that they failed the task but it was “Absolutely Impossible.” Another emotion survey was then filled out. Using a nine-point mood scale, here is the mean decrease in mood once subjects learned of their failure to help Julie:<sup>52</sup>

	Objective Instructions	Empathy-Inducing Instructions
Failure was not Justified	-2.23	-3.17
Failure was Justified	-1.25	-2.83

---

<sup>51</sup> See Batson et al. 1988: 58-59, Batson and Weeks 1996: 148-149, and Batson 2002: 95-96.

<sup>52</sup> Batson and Weeks 1996: 152.

The fact that there was no statistically significant variance to the numbers on the right hand side strongly suggests that subjects were concerned about preventing Julie's suffering rather than about whether they would be punished for not helping her, since if their failure was justified, there would be no grounds for legitimate punishment.<sup>53</sup>

(iii) Empathy-Specific Rewards. The third class of egoistic explanations for empathy-induced helping behavior centers not on punishments but rather on rewards. It claims that people in such cases are ultimately motivated by one or more of the specific rewards attached to helping the person in need with whom they empathize. Such rewards might come in the form of social or religious benefits, or more internal feelings of joy, honor, pride, or pleasure.<sup>54</sup> Once again, if such explanations are correct, then the resulting behavior is not driven by genuine compassion.

One way to test these reward explanations is to examine differences in mood based upon whether the problems of the person in need are relieved by the empathizer or by a third party. If empathy-specific rewards are the motivating force behind helping, then comparatively speaking there should be a much higher mood in those who address the need themselves than in those who observe the need relieved by someone else. According to the empathy-altruism hypothesis, on the other hand, it should not matter who is helping the person in need so long as that need is addressed.<sup>55</sup>

Once again experimental support for this kind of egoistic explanation has not been forthcoming. Batson et al. (1988) devised an experiment in which subjects were initially told that their performance of a helping task would influence the number of electric

---

<sup>53</sup> For additional studies which challenge the empathy-specific punishment hypothesis, see Batson et al. 1988.

<sup>54</sup> For discussion, see Batson et al. 1988, Batson 2002: 96-97, and Batson et al. 2003: 281-282.

<sup>55</sup> See Batson et al. 1988: 53 and Batson 2002: 96-97.

shocks someone else would receive. Similar to the previous study, they then heard a recording of the other person expressing concern about being shocked. Two types of experimental manipulations were then introduced – whether subjects were told that the other person had been reassigned to a task that did not involve receiving shocks or was still assigned to the same task, and whether as a result the subject would still be performing his or her helping task or it was not required anymore. The general idea was that by performing their helping task well, subjects could spare the person from receiving shocks, but if the person is reassigned by those in charge of the experiment, then he would have been prevented from suffering but not through the actions of the subject. Thus if the above egoistic explanation is correct, we would expect significant variations on the mood scores of the subjects after the manipulations were done. Instead, the mean ratings on a mood index from 1 to 9 (with 9 being good mood) were as follows:<sup>56</sup>

	Subject Performs Helping Task	Does Not Perform
Prior Reassignment of the Other Person from Shock Task	6.29	6.73
No Prior Reassignment of the Other Person from Shock Task	6.56	5.84

The overall conclusion was that “subjects’ self-reported mood provided no evidence that high-empathy subjects felt better when the victim’s need was relieved by their own action than when it was relieved by other means.”<sup>57</sup>

Stepping back from these different egoistic accounts of the motive behind helping behavior exhibited by empathetic subjects, the general point is this. According to over thirty different experiments carried out by Batson and his colleagues, there is no evidence

---

<sup>56</sup> Batson et al. 1988: 56.

<sup>57</sup> Ibid.

that high levels of empathy generate egoistic motivation to help. And this should be welcome news to the virtue ethicist, since then we could have in place another important constituent of the virtue of compassion. Although we do not have the relevant longitudinal studies yet, global helping traits certainly appear to be cross-situationally consistent and iteratively stable, and now it seems that they can be triggered by motives which are genuinely altruistic and thus can serve as the motives of a compassionate person. Perhaps then GHTs can become the foundational building blocks for an empirically informed and respectable account of some of the central psychological mechanisms responsible for the instantiation of a virtue like compassion.

## **6. Two Lingering Worries for Virtue Ethics**

With global helping traits and a source for altruistic motivation, perhaps the virtue ethicist now has in hand the tools needed for responding to concerns about compassion arising from social psychology. Nevertheless, there are two remaining reasons for thinking that such concerns will not go away so easily – most people do not seem to exhibit compassionate behavior, and even when they do empathize with others, the feelings which result may be too fragile of a basis for compassion.

*The Extent of Compassion.* As a preliminary point, we should agree with Harman and Doris that the social psychology literature does suggest that most people do not seem to exhibit compassionate behavior when an opportunity to help arises. Recall that a global character trait such as compassion enables the person who possesses it to exhibit cross-situationally consistent trait-relevant behavior. Thus we would expect that if most people had this trait they would be compassionate in a wide variety of situations. But this is not

what we in fact find. Most control subjects would not volunteer to tutor Carol at her home or prevent someone from being mildly shocked.<sup>58</sup> Note as well that the helping tasks in question are often not particularly strenuous ones like making sizable donations to charity or forgoing a lucrative career to care for sick relatives. In such cases we might expect that even moderately compassionate people would not always rise to the challenge. Instead the lack of compassionate behavior that we find in the social psychology literature is manifested with respect to such relatively painless actions as making change, picking up dropped papers, or pledging an hour of time.<sup>59</sup>

To this point about the extent of compassion, it might be responded that the helping actions involved in the relevant psychological studies did not concern people in serious need, and hence such opportunities to help may have simply escaped the notice of a compassionate person who is not experiencing empathy.<sup>60</sup> However, I find this objection to be less than compelling. First of all, the helping tasks were ones for which there were no other helping opportunities in the immediate environment which might have been thought to be more significant than the experimental helping opportunity. Rather, the alternative to performing the helping task was simply to ignore Carol's request for help in the class or allow a person to experience electric shocks about which he was clearly fearful. Secondly, it is hard to accept the claim that the helping opportunities in the relevant experiments were not significant. Recall that subjects were given a chance to help someone who had just broken her legs, or whose parents had died and was suddenly responsible for taking care of two younger siblings. It seems difficult

---

<sup>58</sup> For the latter helping opportunity, see Batson et al. 1983.

<sup>59</sup> For experiments involving these helping tasks, see Baron 1997, Batson et al. 1979, and Weyant 1978 respectively.

<sup>60</sup> I have been presented with this objection on several occasions, but it was stated perhaps most forcefully by James Taylor in written comments on my 2007.

to explain the failure of subjects to perform these actions without the benefit of empathy by asserting that they fall outside the scope of compassionate concern.

Of course, the virtue ethicist who wants to make use of GHTs in order to develop an empirically informed account of compassion could respond by saying that all this absence of compassionate behavior shows is that the subjects' GHTs were not triggered in the right way – there were no feelings of empathy for the person in need, for instance. Once a trigger was activated, we know that other things being equal helping behavior should have increased significantly.<sup>61</sup>

But this response will not do. A genuinely compassionate person would not need to have his or her helping behavior triggered by inputs such as a motive fostered by empathetic feelings, or for that matter by any of the other triggers mentioned in section two such as positive or negative moods or feelings of embarrassment. Even *without* such feelings, a compassionate individual would be expected, other things being equal, to be able to recognize when someone is in need and try to do what he can to help.<sup>62</sup> And we would not count it as a legitimate excuse that the reason why someone was not compassionate in a given instance was that he could not work himself up into a good

---

<sup>61</sup> Alternatively, the virtue ethicist could instead attempt to explain the absence of compassionate behavior in terms of GHTs actually being *inhibited* from activating. However, this seems like a much less promising strategy to adopt. For it is difficult to come up with plausible inhibitors for all the control subjects who did not help in malls, schools, fairs, psychology labs, and other locations where such experiments were conducted. In my view it is far more natural to attempt to explain subjects' failure to help in terms of their GHTs not being activated than by saying that they were actually inhibited.

<sup>62</sup> An anonymous referee has claimed instead that part of what it is to be a virtuous person is to always self-generate empathetic feelings when confronted by someone in need. I find the claim made in the text to be more intuitively correct, and so let me try to briefly make two points in its favor. First, many helping opportunities are immediate, and so require an automatic, non-reflective helping response by the virtuous agent. But as we noted in section three, paradigm cases of empathy involve cognitive acts of imaginatively considering what the other person perceives in the situation and feels as a result. And it is at least not obvious how those acts could be performed automatically and non-reflectively. Secondly, it seems possible to imagine cases in which a compassionate person helps a stranger without first empathizing with him because the stranger is from a radically different culture or has an unfamiliar set of values. In such cases, it might not be clear to the virtuous person how to even go about making sense of what the stranger would be feeling in the situation.

mood, or did not have any empathetic feelings towards the person in need. These excuses would be seen as, quite frankly, rather lame. This, then, is the main explanation for why global helping traits are not equivalent to traits of compassion, and so explains why even if they do exist, GHTs do not automatically vindicate the empirical adequacy of the virtue ethicist's commitment to the existence of compassion.<sup>63</sup>

To be fair, an important concession should be made to virtue ethicists at this point. Even though the psychological data suggests that most people do not exhibit compassion, it does not rule out the possibility that perhaps a few actually do possess such a virtue. Roughly 33% of controls helped Carol in Toi and Batson's experiment without the empathy manipulation, and similarly 29% volunteered to help in the low benefits / high costs scenario of Weyant's experiment. Furthermore, even if, when they feel empathy, such subjects do experience extra motivation to help from their empathetic state, this surplus motivation might simply end up over-determining their helping behavior since we already know that they help others in these situations even without this extra contribution.

If at least some of these subjects really were compassionate, then the fact that there were only a few of them need serve as no embarrassment for virtue ethics. After all, it is no part of the virtue ethical tradition that we should expect there to be widespread possession of any virtue or vice. Emphasis in this tradition has been placed on how the

---

<sup>63</sup> A second reason why GHTs do not amount to compassionate traits is that, while empathy supplies them with a motive which is altruistic, they can also be triggered by self-interested motives as well. In the case of moods, for example, we saw briefly in section two that the leading model of the relationship between negative moods and helping is the mood management hypothesis, according to which a bad mood generates a motive to relieve the bad mood and return the person to an equilibrium condition. A number of means might be available for elevating mood, and one of them will often be helping others because of the social rewards and gratification associated with such behavior. Thus helping is treated merely as a means to benefiting the agent, but such egoistic motivation is incompatible with genuine compassion.

For more on the mood management hypothesis, see Cialdini, Darby, and Vincent 1973, Cialdini and Kenrick 1976, Weyant 1978, Manucia, Baumann, and Cialdini 1984, Cialdini et al 1987, Schaller and Cialdini 1990, and Taylor 1991 as well as my 2007.

life of progression to full virtue is one of continuous struggle in overcoming character defects and external obstacles. Thus for the Plato of *The Republic*, virtue emerges through participation in a long and demanding educational process, while for Aristotle virtues are character traits that must be habituated in children and positively reinforced in adults over extended periods of time.<sup>64</sup>

Thus I am inclined to agree that the experimental results at issue in this paper do not preclude the instantiation of genuine compassion in some of the test subjects, although they do not provide any positive evidence for its being possessed either. And even Doris himself concedes that the social psychology literature is compatible with the claim that “some individuals may quite consistently exhibit compassion.”<sup>65</sup> At the same time, Doris and company do seem to be right when they claim that *most* people do not appear to exhibit behavior which is compassionate. And it is also no longer clear whether GHTs are the proper building blocks for thinking about what makes a few people potentially compassionate, or about how we might try to improve the degree to which everyone else possesses this trait.

*Empathy Induced Motivation is Extremely Fragile.* Suppose that the virtue ethicist is prepared to accept that global helping traits are in fact widely possessed, but denies my claim that the mere fact that they must always be triggered by empathetic feelings in order to lead to altruistic motivation precludes them from serving as the basis for compassion. Then given the fact that so few controls in the relevant experiments seem to actually be exhibiting compassionate behavior, we will need a story about how *more*

---

<sup>64</sup> See in particular *Nicomachean Ethics* 1099b29-32, 1103b16-31, 1152a30-34, 1179b25-29, 1180a1-5, 15-19 and Burnyeat 1980. These points are developed in greater detail in my 2003.

<sup>65</sup> Doris 2002: 65. Similarly, he notes that “situationism does not preclude the existence of a few saints, just as it does not preclude the existence of a few monsters” (60).

people can instantiate the virtue of compassion. And if the virtue ethicist wants to make use of GHTs to develop an empirically informed account of the psychological mechanisms responsible for compassion, then the natural place to look is to see whether people can be disposed to be more empathetic than they typically are on a daily basis, thereby leading to the repeated triggering of their GHTs.<sup>66</sup>

But here the psychological results seem to put up a significant obstacle to such a proposal. For as some of Batson's other work has shown, empathetic feelings that have been induced in subjects are extremely fragile. In other words, subjects can be easily distracted from thinking about what the person in need is feeling, and when they are, their altruistic motivation simply vanishes. Even worse, when made to think about themselves instead, not only will these subjects lose altruistic motivation to help but any remaining motivation will be self-interested.<sup>67</sup>

We can see this in dramatic form by comparing two studies reported in Batson et al. (1983). Subjects were initially supposed to watch a participant perform a series of trials during which he or she would receive electric shocks. In the easy to escape condition, subjects had to only observe two trials, whereas in the difficult to escape condition, they had to watch all of them (between two and ten). However, after the trials

---

<sup>66</sup> Note that I am not simply assuming that because the virtue of compassion and a GHT (when triggered) both produce helping behavior, it follows that they must have a similar mechanism. Rather, I am suggesting that, provided the virtue ethicist denies the claim at the start of the above paragraph, GHTs become a natural place to look for a psychological mechanism upon which to build an empirically respectable account of the virtue of compassion. And the idea here for doing so is that an ordinary GHT might become a crucial component of such a virtue if we can find ways to dispose people to become more naturally empathetic on a regular basis.

Of course, the virtue ethicist is free to not make use of GHTs at all in trying to develop an empirically legitimate account of the virtue of compassion, although it is unclear where else in the social psychology literature she would go for help. Thanks to an anonymous referee for raising these issues.

<sup>67</sup> Batson 1987: 109.

began, subjects were given a chance to switch places with the participant and receive the shocks themselves. Here is the crucial part of what they were told:

Before you decide, I should tell you that the shocks Elaine (Charlie) has been getting, and that you would receive if you took her (his) place, are Level 1 shocks. They're the lowest level of shock that would be perceived as at all aversive.<sup>68</sup>

The percentage who agreed to help and the mean number of shock trials agreed to be taken were as follows:<sup>69</sup>

	Distress	Empathy
Easy Escape	25% (1.25)	86% (5)
Difficult Escape	89% (4.89)	63% (3.75)

In the second study, the procedure was the same except that the passage above read:

Before you decide, I should tell you that the shocks Elaine (Charlie) has been getting, and that you would receive if you took her (his) place, are Level 4 shocks. They are clearly painful, but of course not harmful.<sup>70</sup>

This one change led to a stunning difference in the results:<sup>71</sup>

	Distress	Empathy
Easy Escape	50% (3.13)	14% (.29)
Difficult Escape	67% (3.89)	60% (3.30)

Compassionate people, on the other hand, would not let their attention to the needs of others be distracted so easily by threats to their own self-concerns, or at least not unless the treats were much more severe.

Of course, the virtue ethicist could just argue that people need to be better habituated into blocking out other considerations from their minds and focusing solely on what the person in need is imagined to be feeling. But there are real practical limitations

---

<sup>68</sup> Batson et al. 1983: 714.

<sup>69</sup> Ibid.

<sup>70</sup> Ibid., 715.

<sup>71</sup> Ibid., 716.

to how much the psychological lives of most people can be reshaped, and the worry here is that it may be asking too much of people to expect that they both develop a robust disposition to be empathetic on a daily basis and also have their empathetic emotions, once triggered, be routinely strong enough to preclude the intruding influences of self-concern.<sup>72</sup>

## 7. Conclusion

Given these two pieces of bad news for the virtue ethicist – that few people seem to exhibit compassionate behavior when an opportunity to help arises, and that even empathy induced motivation to help is extremely fragile – it might seem as if this paper ultimately just ends up providing further support for situationism, rather than defending an intermediate position between situationism and traditional virtue ethics. However, situationists are committed to rejecting not just the existence of familiar virtues like compassion and honesty, but rather the existence of *all* global character traits as such. And while the two pieces of news in the previous section of this paper might serve as bad news for virtue ethicists, they have no bearing on the central claim of this paper, namely that the experimental results we have examined on empathy and helping behavior are compatible with the existence of global helping traits. Thus the (widespread) existence of

---

<sup>72</sup> A related literature that might be helpful to examine in considering these issues further is the literature on stereotype activation and control. Devine and Monteith 1999, for example, argue that while it is difficult, psychological reshaping in order to either inhibit or override stereotype activation has been experimentally supported in a range of psychological work. Bargh 1999, on the other hand, has argued for a much less optimistic view.

Given limitations of space, I cannot pursue what interesting connections there might be here to the reshaping needed in order to make empathetic subjects not deviate from focusing solely on what the person in need is experiencing. I hope to examine such connections in future work, and so admit that for the time being there might be a fruitful response in this neighborhood for the virtue ethicist to use in order to address the worry raised in the text above. I am grateful to Nancy Snow for helpful comments on these issues.

a traditional character trait like compassion might be in trouble, but it is premature to conclude from such a claim that no global traits whatsoever exist in this area.

Finally, it is important to note that this paper is not intended to serve as an exhaustive survey of the empirical literature on empathy and helping, nor as a detailed treatment of all the implications that literature might have for the current debate between virtue ethicists on the one hand and Harman and Doris on the other. Rather it is primarily intended to introduce a new body of social psychology research to philosophers working in this area, and to show how that research can contribute to carving out an intermediate position between the two opposing sides. Much further discussion will be needed in order to see whether this proposal concerning the existence and nature of global helping traits can survive as a viable approach.<sup>73</sup>

---

<sup>73</sup> For very helpful written comments, I am grateful to Nancy Snow and several anonymous referees. This paper is a companion piece to my 2007, which focuses on helping in relation to good and bad moods rather than empathy.

## Works Cited

- Apsler, R. (1975). "Effects of Embarrassment on Behavior Toward Others." *Journal of Personality and Social Psychology* 32: 145-153.
- Aristotle. (1985). *Nicomachean Ethics*. Trans. Terence Irwin. Indianapolis: Hackett Publishing Company.
- Athanassoulis, Nafsika. (2001). "A Response to Harman: Virtue Ethics and Character Traits." *Proceedings of the Aristotelian Society* 100: 215-222.
- Bargh, John. (1999). "The Cognitive Monster: The Case Against the Controllability of Automatic Stereotype Effects," in *Dual-Process Theories in Social Psychology*. Ed. S. Chaiken and Y. Trope. New York: Guilford Press. 361-382.
- Baron, Robert. (1997). "The Sweet Smell of . . . Helping: Effects of Pleasant Ambient Fragrance on Prosocial Behavior in Shopping Malls." *Personality and Social Psychology Bulletin* 23: 498-503.
- Batson, C. (1987). "Prosocial Motivation: Is it Ever Truly Altruistic?" in *Advances in Experimental Social Psychology*. Vol. 20. Ed. L. Berkowitz. San Diego: Academic Press, 65-122.
- \_\_\_\_\_. (1991). *The Altruism Question: Toward a Social-Psychological Answer*. Hillsdale: Erlbaum.
- \_\_\_\_\_. (2002). "Addressing the Altruism Question Experimentally," in *Altruism & Altruistic Love: Science, Philosophy, & Religion in Dialogue*. Ed. S. Post, L. Underwood, J. Schloss, and W. Hurlbut. Oxford: Oxford University Press. 89-105.
- Batson, C., J. Coke, F. Chard, D. Smith, and A. Taliaferro. (1979). "Generality of the 'Glow of Goodwill': Effects of Mood on Helping and Information Acquisition." *Social Psychology Quarterly* 42: 176-179.
- Batson, C., B. Duncan, P. Ackerman, T. Buckley, and K. Birch. (1981). "Is Empathic Emotion a Source of Altruistic Motivation?" *Journal of Personality and Social Psychology* 40: 290-302.
- Batson, C., K. O'Quin, J. Fultz, M. Vanderplas, and A. Isen. (1983). "Influence of Self-Reported Distress and Empathy on Egoistic Versus Altruistic Motivation to Help." *Journal of Personality and Social Psychology* 45: 706-718.
- Batson, C., J. Dyck, J. Brandt, J. Batson, A. Powell, M. McMaster, and C. Griffitt. (1988). "Five Studies Testing Two New Egoistic Alternatives to the Empathy-Altruism Hypothesis." *Journal of Personality and Social Psychology* 55: 52-77.
- Batson, C. J. Batson, C. Griffitt, S. Barrientos, J. Brandt, P. Sprengelmeyer, and M. Bayly. (1989). "Negative-State Relief and the Empathy-Altruism Hypothesis." *Journal of Personality and Social Psychology* 56: 922-933.
- Batson, C. and Joy Weeks. (1996). "Mood Effects of Unsuccessful Helping: Another Test of the Empathy-Altruism Hypothesis." *Personality and Social Psychology Bulletin* 22: 148-157.
- Batson, C., S. Early, and G. Salvarani. (1997). "Perspective Taking: Imagining How Another Feels Versus Imagining How You Would Feel." *Personality and Social Psychology* 23: 751-758.
- Batson, C. 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: 495-509.
- Batson, C., P. van Lange, N. Ahmad, and D. Lishner. (2003). "Altruism and Helping Behavior," in *The Sage Handbook of Social Psychology*. Ed. M. Hogg and J. Cooper. London: Sage Publications, 279-295.

- Burnyeat, M. (1980). "Aristotle on Learning to Be Good," in *Essays on Aristotle's Ethics*. Ed. Amélie Rorty. Berkeley: University of California Press. 69-92.
- Campbell, John. (1999). "Can Philosophical Accounts of Altruism Accommodate Experimental Data on Helping Behavior?" *Australasian Journal of Philosophy* 77: 26-45.
- Carlson, M., V. Charlin, and N. Miller. (1988). "Positive Mood and Helping Behavior: A Test of Six Hypotheses." *Journal of Personality and Social Psychology* 55: 211-229.
- Cialdini, R., B. Darby, and J. Vincent. (1973). "Transgression and Altruism: A Case for Hedonism." *Journal of Experimental Social Psychology* 9: 502-516.
- Cialdini, R. and D. Kenrick. (1976). "Altruism as Hedonism: A Social Development Perspective on the Relationship of Negative Mood State and Helping." *Journal of Personality and Social Psychology* 34: 907-914.
- Cialdini, R. D. Baumann, and D. Kenrick. (1981). "Insights from Sadness: A Three-Step Model of the Development of Altruism as Hedonism." *Developmental Review* 1: 207-223.
- Cialdini, R., M. Schaller, D. Houlihan, K. Arps, J. Fultz, and A. Beaman. (1987). "Empathy-Based Helping: Is It Selflessly or Selfishly Motivated?" *Journal of Personality and Social Psychology* 52: 749-758.
- Cialdini, R., S. Brown, B. Lewis, C. Luce, and S. Neuberg. (1997). "Reinterpreting the Empathy-Altruism Relationship: When One Into One Equals Oneness." *Journal of Personality and Social Psychology* 73: 481-494.
- Cialdini, R., and J. Fultz. (1990). "Interpreting the Negative Mood-Helping Literature via 'Mega'-Analysis: A Contrary View." *Psychological Bulletin* 107: 210-214.
- Coke, J. C. Batson, and K. McDavis. (1978). "Empathic Mediation of Helping: A Two-Stage Model." *Journal of Personality and Social Psychology* 36: 752-766.
- Darley, J. and C. Batson. (1973). "'From Jerusalem to Jericho:' A Study of Situational and Dispositional Variables in Helping Behavior." *Journal of Personality and Social Psychology* 27: 100-108.
- Darwall, Stephen. (1998). "Empathy, Sympathy, Care." *Philosophical Studies* 89: 261-282.
- DePaul, Michael. (1999). "Character Traits, Virtues, and Vices: Are There None?" *Proceedings of the World Congress of Philosophy*. Philosophy Documentation Center. 1: 141-157.
- Devine, Patricia and Margo Monteith. (1999). "Automaticity and Control in Stereotyping," in *Dual-Process Theories in Social Psychology*. Ed. S. Chaiken and Y. Trope. New York: Guilford Press. 339-360.
- Doris, John. (1998). "Persons, Situations, and Virtue Ethics." *Noûs* 32: 504-530.
- \_\_\_\_\_. (2003). *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press.
- Dovidio, John and Louis Penner. (2001). "Helping and Altruism," in *Blackwell Handbook of Social Psychology: Interpersonal Processes*. Ed. G. Fletcher and M. Clark. Oxford: Blackwell, 162-195.
- Eisenberg, Nancy and Paul Miller (1987). "Empathy, Sympathy, and Altruism," in Eisenberg and Strayer 1987a: 292-316.
- Eisenberg, Nancy and Janet Strayer. (1987a). *Empathy and Its Development*. Cambridge: Cambridge University Press.
- \_\_\_\_\_. (1987b). "Critical Issues in the Study of Empathy," in Eisenberg and Strayer 1987a, 3-13.
- Harman, Gilbert. (1999). "Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error." *Proceedings of the Aristotelian Society* 99: 315-332
- \_\_\_\_\_. (2000). "The Nonexistence of Character Traits." *Proceedings of the Aristotelian Society* 100: 223-226.
- Isen, A. (1987). "Positive Affect, Cognitive Processes, and Social Behavior," in *Advances in Experimental Social Psychology*. Ed. L. Berkowitz. San Diego: Academic Press. 203-254.

- Latané, Bibb and J. M. Darley. (1970). *The Unresponsive Bystander: Why Doesn't He Help*. New York: Appleton-Century-Crofts.
- Latané, Bibb and Steve Nida. (1981). "Ten Years of Research on Group Size and Helping." *Psychological Bulletin* 89: 308-324.
- Kamtekar, Rachana. (2004). "Situationism and Virtue Ethics on the Content of Our Character." *Ethics* 114: 458-491.
- Manucia, G., D. Baumann, and R. Cialdini. (1984). "Mood Influences on Helping: Direct Effects or Side Effects?" *Journal of Personality and Social Psychology* 46: 357-364.
- Milgram, Stanley. (1963). "Behavioral Study of Obedience." *Journal of Abnormal and Social Psychology* 67: 371-378.
- Miller, Christian. (2003). "Social Psychology and Virtue Ethics." *The Journal of Ethics* 7: 365-392.
- \_\_\_\_\_. (2007). "Social Psychology, Mood, and Helping Behavior: Mixed Results for Virtue Ethics." Under Review.
- Oliner, S and P. Oliner. (1988). *The Altruistic Personality: Rescuers of Jews in Nazi Europe*. New York: Free Press.
- Regan, J. (1971). "Guilt, Perceived Injustice, and Altruistic Behavior." *Journal of Personality and Social Psychology* 18: 124-132.
- Sabini, John and Maury Silver. (2005). "Lack of Character? Situationism Critiqued." *Ethics* 115: 535-562.
- Schaller, M. and R. Cialdini (1990). "Happiness, Sadness, and Helping: A Motivational Integration," in *Handbook of Motivation and Cognition*. Ed. E. Higgins and R. Sorrentino. New York: The Guilford Press. 265-296.
- Slote, Michael. (2004). "Autonomy and Empathy." *Social Philosophy and Policy* 21: 293-309.
- Snow, Nancy. (2000). "Empathy." *American Philosophical Quarterly* 37: 65-78.
- Sober, Elliott and David Sloan Wilson. (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge: Harvard University Press.
- Sreenivasan, Gopal. (2002). "Errors about Errors: Virtue Theory and Trait Attribution." *Mind* 111: 47-68.
- Stich, Stephen, J. Doris, and E. Roedder. (2007). "The Science of Altruism." Manuscript.
- Taylor, Shelley. (1991). "Asymmetrical Effects of Positive and Negative Events: The Mobilization-Minimization Hypothesis." *Psychological Bulletin* 110: 67-85.
- Toi, M. and C. Batson. (1982). "More Evidence that Empathy is a Source of Altruistic Motivation." *Journal of Personality and Social Psychology* 43: 281-292.
- Weyant, J. (1978). "Effects of Mood States, Costs, and Benefits on Helping." *Journal of Personality and Social Psychology* 36: 1169-1176.
- Wispé, Lauren. (1986). "The Distinction between Sympathy and Empathy: To Call Forth a Concept, A Word is Needed." *Journal of Personality and Social Psychology* 50: 314-321.
- Zimbardo, P. G., W. Banks, C. Haney, and D. Jaffee. (1973). "The Mind is a Formidable Jailer: A Pirandellian Prison." *New York Times Magazine*. April 8.