Virtue, Vice, and Situationism

Tom Bates and Pauline Kleingeld

University of Groningen

Forthcoming in:


**Abstract**

On the basis of psychological research, a group of philosophers known as 'situationists' argue that the evidence belies the existence of broad and stable (or 'global') character traits. They argue that this condemns as psychologically unrealistic those traditions in moral theory in which global virtues are upheld as ideals. After a survey of the debate to date, this article argues that the thesis of situationism is ill-supported by the available evidence. Situationists overlook the explanatory potential of a large class of global character traits, namely, vices that do not involve other-directed malevolence, such as laziness, cowardice, and selfishness. A detailed discussion of the relevant empirical studies bearing on moral psychology shows that the behavioral patterns observed in these studies are consistent with the widespread possession of such non-malicious vices. This means, contrary to the situationist thesis, that the empirical record is fully compatible with the common existence of global character traits.

**Keywords**: character, character trait, behavioral consistency, moral psychology, situationism, vice, virtue.

**Introduction**

Virtues are usually regarded as character traits. Character traits are regarded as dispositions to reason, feel, and act in trait-appropriate ways across a variety of trait-relevant situations. That is to say, character traits are associated with consistent patterns of behavior. They are usually conceived as 'global' or 'robust' in the sense that they are not tied to one specific situational context but supposed to manifest in trait-appropriate behavior across a variety of trait-relevant situations. Global character traits are usually conceived as stable, in the sense that the associated behavior is not easily disrupted by morally irrelevant situational variables. A global or robust trait of bravery, for example, will manifest when bravery is called for, whether on the battlefield, on the rollercoaster, or in the face of morally inappropriate requests of one's superiors, and regardless of whether one is strongly
encouraged by one's friends or all alone. This conception of character traits is central to Aristotelian virtue ethics, but insofar as virtues are firm and stable dispositions to do what is morally required, they are granted crucial importance by many non-Aristotelians as well.

The existence of global character traits, as a matter of ordinary human psychology, has been called into question on empirical grounds, however, by several philosophers known as 'situationists'. John Doris and Gilbert Harman, among others, claim that empirical research in social psychology shows that global character traits hardly exist (e.g., Doris 1998, 2002; Harman 1999, 2000, 2009; Merritt et al. 2010; Vranas 2005). Systematic observation in experimental settings indicates, they argue, that morally irrelevant or insignificant situational variables have huge effects on people's morally relevant behavior. For example, the vast majority of people believes that one ought not inflict harm on others; yet, as Stanley Milgram shows in a famous series of experiments, a large majority of ordinary people are willing to deliver extremely painful and even fatal electric shocks to a likeable man who makes mistakes in a memory test, at the mere request of an experimenter and despite the fact that the victim is screaming in pain and has withdrawn his consent to the experiment (Milgram 1974). Other studies show that the presence of inactive bystanders strongly reduces the rate at which people help others in need, and that being in a neutral mood, as compared to a good or bad mood, also reduces helping rates. Yet most people deny that, other things being equal, the mere presence of inactive bystanders, or being in a neutral mood as compared to a good one, are valid reasons not to help someone in need. If people generally did possess global character traits such as helpfulness or kindness, situationists argue, one would expect the subjects in these studies to behave rather differently. Given how easy it is to influence behavioral patterns in test subjects by introducing morally irrelevant or insignificant situational variables, the attribution of global character traits is generally unwarranted. Situationists conclude that behavior consistent with global traits is 'rare enough to count as
abnormal' (Doris 2002, 65). In their view, behavior is typically best explained by reference to situational factors that trigger subconscious and 'depersonalized' response-tendencies that are largely independent of agents' moral values (Merritt et al. 2010, 370). Although some situationists are more radical in their rejection of 'character' than others, they all agree that few, if any, people possess global character traits.

Merritt, Doris and Harman formulate the core argument as the following modus tollens:

1. If behavior is typically ordered by robust traits, systematic observation will reveal pervasive behavioral consistency.
2. Systematic observation does not reveal pervasive behavioral consistency.¹
3. Therefore, behavior is not typically ordered by robust traits. (Merritt et al. 2010, 357-8, identical in Doris 2005, 633)

Situationists argue that their thesis spells trouble for Aristotelianism in particular. For if robust (or 'global') traits hardly exist, this seems to condemn Aristotelian virtue ethics as unrealistic, given that such traits are central to its conception of ethics (Alfano 2013; Doris 1998, 2002; Harman 1999, 2000, 2009; Merritt 2000; Vranas 2005, Merritt et al. 2010). Their challenge has prompted a heated debate in ethics and moral psychology.

In this essay, we argue that the situationists' core thesis is empirically ill-supported. Situationists fail adequately to consider the explanatory potential of a key class of global character traits. They do consider one set of vices, namely, those involving dispositions to harm others, such as cruelty. But they fail to consider the possibility that much human behavior stems from global vices that are not associated with the pursuit of harm to others for its own sake, such as selfishness, cowardice, or laziness (in the morally objectionable sense). If one takes the possibility of this broader range of vices seriously, the empirical evidence suggests that global character traits may well be abundant.
In Section 1, we survey the debate to date and set the stage for the argument in the sections to follow. In Section 2, we analyze the structure of the reasoning in support of the thesis of situationism. We locate the flaw that makes the core argument unsound, and we discuss the role of vices in the assessment of the empirical evidence. In Section 3, we show that the very empirical evidence that situationists regard as the key support for their thesis in fact admits of an alternative explanation in terms of global traits. We conclude that the available evidence does not count in favor of situationism.

In arguing for this conclusion, we do not mean to deny the important influence of situational factors on human behavior, of course. All sides in the debate agree that situational variables affect behavior and that human behavior often reliably varies in response to variations in the circumstances (to give a trivial example: if it starts raining, drivers tend to turn on their windshield wipers). What is at issue in the debate we shall be discussing in this essay is the fact that morally relevant behavior very often reliably covaries with morally insignificant differences in situational circumstances. The question is whether such patterns can plausibly be understood as stemming from the agents' global character traits, that is, as reflecting the agents' sensitivity to trait-relevant differences between the situations.

Situationists deny this, arguing that such patterns are best explained by reference to the situational variables and the automatic response tendencies they trigger. We argue that their argument for this thesis is flawed because they overlook a possible trait-based explanation of the patterns.

1. Situationism and Virtue Ethics: The State of the Debate

Situationists argue that Aristotelian virtue ethicists ignore or fail to realize that the stipulated connection between inner dispositions and observable behavioral patterns, which is crucial to the idea of virtuous character, is not confirmed empirically (Doris 2002, 15-22). If
such a connection does not exist, then the virtue ethical ideal that is built on the assumption of such a connection turns out to be unrealistic. Most of the ensuing debate has focused on whether situationism really does pose a problem for Aristotelian virtue ethics. Many critics of situationism believe that the attack can be deflected, but situationists are not convinced by the rebuttals, and it seems that the debate has reached a stalemate. We start with a brief survey of the main argumentative moves concerning three central issues.

First, several authors have replied that the empirical evidence on which situationists rely does not actually show the nonexistence of global character traits (Miller 2003, 2009, 2010; Snow 2010; Wielenberg 2006). At most it simply shows that virtues are rare—and the data show that a non-negligible minority of subjects do display ‘good’ behavior. The rarity of virtue is not a challenge to Aristotelian virtue ethics—so the reply goes—because Aristotelians do not typically claim that virtuous character is widespread (e.g., Kamtekar 2004, 482–485). After all, on the Aristotelian conception, virtue in the full sense requires practical wisdom and the alignment of one’s feelings with one’s moral insight, and this is a demanding ideal. Equally importantly, genuine virtue is impossible to detect with the methods of the psychological experiments on which situationists base their argument.

Different agents may perform the same observable behavior while doing so from radically different motives (Kristjánsson 2008). Other authors (Russell 2009; Snow 2010; Webber 2013) have argued that the empirical evidence in support of the ‘cognitive-affective personality system’ (CAPS) theory of personality traits, as developed by Mischel and Shoda (1995), can be used to support character traits, if traits are understood in terms of that theory.

Second, and relatedly, critics have argued that situationists work with a mistaken conception of character. Situationists have been criticized for using a ‘behaviorist’ notion of traits (Webber 2006, 2007), for reducing virtues to ‘stereotypical behavior, in isolation from how people reason’ (Kamtekar 2004, 460, 477), and for failing to recognize the intellectual
nature of character (Annas 2005, 639; Kamtekar 2004, Kristjánsson 2008, Kupperman 2001, Sreenivasan 2002). It is clear why situationists can be interpreted this way: they state that the criterion for appropriate trait attribution is that the trait be ‘reliably manifested’ under the appropriate ‘trait-relevant eliciting conditions’ (Doris 2002, 22). This makes it seem as if having a virtuous trait quasi-mechanically results in the appropriate response under the relevant circumstances. Instead, critics emphasize, character is an internal disposition, and one can possess a certain character trait even if one does not display the corresponding behavior on one particular occasion. For example, agents may not understand situations in the same way as the observer; hence what may look like inconsistency to an observer may be consistent when understood in terms of the agent's construal of the situation. Critics have also argued that situationists fail to take account of the fact that character is typically conceived as being composed of many traits, and that in any given situation, different traits may pull in different directions. If the agent ends up acting on the basis of one trait, this does not mean she does not possess the others (Webber 2006, Kamtekar 2004). Also, an agent may possess a trait but not in full (Miller 2003, 378-379) or have some traits but not others (Miller 2010).

Third, critics of situationism have argued that there are alternative trait-based explanations for the empirical evidence. John Sabini and Maury Silver, for example, have argued that some of the experimental findings on which situationists rely can be explained in terms of subjects wanting to avoid embarrassment. Neera Badhwar has proposed an explanation of the Milgram results in terms of the subjects’ ‘pusillanimity’. These alternative explanations are said to indicate that the scope of the situationists’ argument is more limited than they claim (Sabini and Silver 2005; Badhwar 2009, 278-287).

Virtue ethicists often regard these rebuttals as sufficient to put the matter to rest, but from the situationists' perspective, there are rejoinders to these criticisms (see Doris 2002; most extensively Alfano 2013, ch.3; on the dialectic of the debate, see also Prinz 2009, 120-
First, when critics of situationism point out that character could simply be rare, situationists respond that their thesis still holds for the vast majority of cases. They regard ‘very rare’ as bad enough (Doris 2002, 115, where this objection is anticipated). More importantly, however, situationists insist that their challenge should not be reduced to the claim that virtuous behavior is rare. Their point is rather that morally relevant behavior typically varies with morally irrelevant or insignificant features of the situation, rather than with any alleged global traits of agents. This is the real challenge to virtue ethics (Doris 2002, 35). As Merritt, Doris and Harman put it: ‘It is not that people fail standards for good conduct, but that people can be induced to do so with such ease’ (Merritt et al. 2010, 357). The situationists’ point is that the factors that best explain behavior are typically not the agents' alleged character traits, but, rather, largely morally irrelevant or insubstantial features of situations that trigger largely unconscious responses in broad segments of the population, regardless of their individual values and beliefs.

Second, situationists deny that their conception of traits is behaviorist or unintellectual. They have always acknowledged that Aristotelian virtues are ‘not mere dispositions but intelligent dispositions, characterized by distinctive patterns of emotional response, deliberation, and decision as well as by more overt behavior’ (Doris 2002, 17, emphasis in original). Certainly, in their view, trait-based action, should it exist, does not happen mechanically but stems from the agent’s own ‘evaluative commitments’ (Doris 2002, 20). Their point is, rather, that empirical evidence shows that behavior typically co-varies with morally extraneous differences between experimental conditions, rather than with agents’ alleged inner characteristics, and that the agents’ inner evaluative commitments are therefore apparently largely behaviorally irrelevant (Doris 2002, 16-27). More recently, situationists have emphasized that cognitive processes, too, depend on situational contexts (Merritt et al. 2010, 359-360). Thus, the appeal to the internal nature of dispositions fails to sway
situationists (Merritt et al. 2010, 358-360, 366). In fact, situationists tend to regard the emphasis on the internal nature of traits as a concession of the behavioral irrelevance of character thus conceived. The same is true of the claim that subjects might possess global traits even if these do not manifest in their behavior.

Finally, concerning the third rebuttal, when critics point to reasons that can explain test subjects’ behavior in specific cases, such as fear of embarrassment, the situationists reply that this applies at best to a limited set of studies (Merritt et al. 2010, 367-370):

These explanations [in terms of reasons] suppose that the actor acted on what he took (or on reflection would take) to be a reason, but a large body of empirical work indicates that this may relatively seldom be the case. (Merritt et al. 2010, 369-370) They emphasize that there are many cases in which the discrepancy between the agents’ values and their behavior is caused by different kinds of factors, such as the weather or the noise of a lawnmower. The situational factors are so disparate and so often unconscious or morally insignificant, situationists claim, that it is not possible to explain the evidence in terms of any one specific reason or trait.

2. Why the Core Argument of Situationism Is Unsound

In order to show that the core argument in support of situationism is unsound, we examine more closely just how exactly the evidence is supposed to show that behavior typically cannot be explained by underlying global traits. The broad idea of the situationists' core argument is quite clear: if global traits exist, then, in certain controlled circumstances, they reliably produce the behavior that one expects from agents who possess such traits; so if the expected behavior does not occur under these circumstances then the agents apparently do not have such traits. Of course, this type of argument needs to be made on the basis of a sufficiently large evidential basis of the right kind. Global traits are not expected to produce
trait-manifesting behavior in every trait-relevant situation. If a person refuses to help a thief in a robbery, this is no proof that she is not a helpful person; there may be other moral considerations that make it inappropriate to offer help. But situationists argue that the general absence of the behavior that we would expect to see if global traits did exist, especially under experimental circumstances where the behavior would clearly be expected of someone who possesses the global trait, does give us reason to deny the existence of such traits, at least as a matter of ordinary moral psychology.

What, then, is the behavior that we should expect, according to situationists, if global traits are widespread? Doris’ standard for trait attribution is that trait-consistent behavior is performed over a run of trait-relevant situations, some of which are ‘less than optimally conducive to that behavior’ (Doris 2002, 19). Such situations are diagnostic: they are unfavorable enough to a specific type of behavior that if this type of behavior does occur it is better explained by reference to the corresponding global character trait than by reference to situational factors. He writes, ‘we are justified in inferring the existence of an Aristotelian personality structure when a person’s behavior reliably conforms to the patterns expected on postulation of that structure’ (Doris 1998, 507).

This leads to the question which patterns one should expect if global traits did exist. Doris suggests the following conditional:

If a person possesses a trait, that person will engage in trait-relevant behaviors in trait-relevant eliciting conditions with markedly above chance probability $p$. (Doris 2002, 19, emphasis in original)

One problem here is that psychological experiments generally are not testing for the attribution of specific traits to specific individuals. Subjects are not asked to complete a range of experiments to see if patterns of behavior emerge in individual cases. Rather, the data are usually gathered on the basis of experiments involving many subjects, and the results are
aggregated for the group. Consequently, the situationists reach their conclusion via a further move. They imagine what pattern of behavior would be expected to emerge from these experiments given a substantial number of people possessing specific global traits, and then they compare that pattern with the observed behavior of the group. In other words, the situationists make certain assumptions about the kinds of global traits that would be shared across the population, if global traits did exist. They argue that the observed behavior does not fit a pattern that is consistent with these traits, and on this basis they deny the existence of global traits.

There are two types of traits for which situationists examine the relation between expected and actual behavior, namely, virtuous global traits such as kindness or compassion, and vicious global traits such as cruelty or aggression. Many experiments look at helping behavior and seem to reveal that people often do not help in situations in which the need is obvious and the cost of helping is low. Moreover, it turns out to be remarkably easy to produce significant changes in helping behavior by introducing morally insignificant situational variables. If many people possess a virtue such as compassion, situationists assume, we would expect to see consistent rates of helping across situations that differ only in helping-irrelevant respects. Alternatively, if many people had a vicious global trait, such as cruelty, one would expect them consistently to engage in cruel behavior across situations. The observed behavior does not follow either pattern, however, and on this basis the situationists conclude that people typically do not possess global traits.

The problem with this argument, however, is that the evidence on which situationists rely at best problematizes the widespread existence of the two classes of global traits they consider, namely, virtues and vices that involve malice. In order for their argument to rule out the widespread existence of all types of global character traits, however, situationists should
also rule out another class of global character traits, namely, *non-malicious vices*. This they neglect to do—or so we shall argue.3

Situationists fail to address the possible explanatory value of global vices that do not involve the pursuit of harm to others, such as dispositions to selfishness, cowardice, laziness, and so on. Their standard for the required 'behavioral consistency' is whether people behave in ways that are consistently morally good or consistently morally bad. Yet the vices that they overlook should not be expected to follow this pattern. These vices do not involve the pursuit of what is morally right, and so they should not be expected to manifest in behavior that is consistently morally good. But these traits do not necessarily lead to morally bad behavior, either. The selfish person may do the right thing when doing so is in her interest, the lazy person may do so when it is easy, and the coward when it is safe. Unlike sadists, these agents do not pursue harm to others as such. But they may well cause harm when that which is morally required runs counter to their interests, or when it involves effort or danger.

In other words, in the case of global non-malicious vices, we should expect a pattern of behavior that is neither consistently morally good nor consistently morally bad, but consistently in keeping with the specific vices in question. In order to be able to rule out the widespread existence of such global traits, situationists should test for patterns of moral laziness, cowardice, selfishness, and so on. Instead, however, they infer the general absence of trait-dependent behavioral patterns from the fact that people are neither consistently morally bad nor consistently morally good. This inference is invalid, however, because it neglects the alternative possibility that many people possess global non-malicious vices. In the next section, we show that the experimental evidence situationists appeal to is in fact consistent with the widespread existence of such global vices.
3. The Case of the Missing Vices

In this section, we examine three sets of empirical studies. First, we examine the evidence on which situationists build their case, and we argue that there are possible (and *prima facie* not implausible) explanations of the behavioral patterns in terms of non-malicious vices (3.1). Second, we show that there are some data that lend initial plausibility to the idea that many people possess non-malicious vices (3.2). Third, we consider evidence that might be taken to tell directly against the existence of such vices--namely, experiments in which subjects behave admirably (3.3). We shall not be arguing that the evidence proves the existence of global non-malicious vices, because there is not enough empirical evidence of the right kind to warrant this stronger claim. Rather, we argue that situationists fail to make the empirical case for their view.

3.1. The Evidence Allegedly Supporting Situationism

3.1.1. Milgram's Obedience Experiments

To show that situationists have not given suitable weight to the possibility of non-malicious vices, let us start by examining the way they rule out vice as an explanation of the test subjects' behavior in Milgram’s 'obedience experiments' (Milgram 1974). In a series of related experiments, Milgram showed that ordinary test subjects were willing to obey an experimenter who requested that they administer electric shocks to a 'learner' (who was a confederate presented as a likeable fellow test subject). The experiment was said to examine the influence of punishment on learning, and the test subjects were instructed to administer electric shocks of gradually increasing severity each time the learner made a mistake on a word-pair memory test. Two thirds of the test subjects turned out to be willing to administer extremely painful and even lethal shocks, by pushing levers labelled 'danger: severe shock' (at 375 volts) or 'XXX' (at 435 and 450 volts), despite the fact that the 'learner' explicitly
withdrew his consent at 150 volts and was screaming in pain (except at the highest voltages, when he was silent).

Situationists repeatedly refer to these experiments and consider them 'powerful evidence for situationism' (Doris 2002, 39). People generally regard the behavior of the Milgram subjects as morally wrong, but nevertheless large numbers of test subjects fail to behave in a way that would be consistent with the corresponding virtuous character trait. Moreover, we should not assume that people lie about regarding the behavior as wrong and enjoy the opportunity to hurt the victim. Milgram considered the possibility that the test subjects' behavior was the product of 'deeply aggressive instincts' (Milgram 1974, 71). To test this hypothesis, he ran a variation of the experiment in which the subjects were free to choose the level of the shock they administered. While one subject went to the maximum, and one to 375 volts, all other subjects stopped before 150 volts, with the mean final shock level between 75 and 90 volts (Milgram 1974, 61). The situationists take this to show that the subjects’ willingness to shock the victim did not stem from a desire to hurt him. On this basis they infer that global vice is not the best explanation of the observed behavior. Doris writes that the evidence 'does not suggest that Milgram had stumbled onto an aberrant pocket of sadists', and that it instead proves 'the power of the situation' (Doris 2002, 42). Harman similarly rejects the suggestion that 'extreme personal dispositions are at fault' (Harman 1999, 322) and infers from this that situational variables explain the behavior. Situationists generally regard this as sufficient to rule out vice as an explanation.

But this is too quick. The situationists overlook the set of global traits that do not consist in dispositions to harm others for its own sake but that, morally speaking, can nevertheless be considered to be global vices, that is, as dispositions to act in ways that involve specific forms of moral failure. For example, one could explain the subjects' behavior in terms of cowardice. One could say that they lacked the courage of their convictions, that
they were cowed by the authority of the experimenter. Alternatively, one could explain the subjects' behavior in terms of a disposition to shift responsibility for one's actions to others, in this case to the experimenter (as many subjects explicitly did) or to the victim (to whom some subjects shifted the blame on the grounds that he gave the wrong answers) (Milgram 1974, 203-204). These dispositions are compatible with the observed behavior. There is no need to settle on any one specific disposition to explain the behavior of all or most of the obedient subjects; the observed behavior of different people may have to be explained in terms of different vices. As mentioned above, Milgram's experiments do not provide a sufficient basis on which to establish that specific individuals have specific traits, because doing so would require longitudinal studies of individual subjects. For the purpose of this essay, however, the fact that the data are compatible with alternative explanations in terms of global vices means that systematic observation does reveal behavioral patterns that are consistent with global character traits, contrary to the situationists' claim.

3.1.2. The Group Effect

Other empirical evidence that situationists mention in support of their thesis are studies of group effects. The evidence from such studies shows that helping behavior is significantly reduced in the presence of others. As Doris points out, 'mild social pressures can result in neglect of apparently serious ethical demands' (Doris 2002, 33). He argues that this evidence 'presses charges of empirical inadequacy against characterological moral psychology' because the group effect shows that dispositions are not robust (Doris 2002, 28).

Substantial evidence for the proposition that the presence of other people serves to inhibit the impulse to help was provided by Bibb Latané and John Darley (1970, 38). They describe three processes that might inhibit helping. First, the agent who intervenes risks embarrassment if the situation turns out not to be one that needed their intervention, and the
greater the 'audience' of bystanders is, the higher the cost of unnecessary intervention. Second, agents may look to our peers to help define an apparently ambiguous situation. Their inaction may lead agents to believe that inaction is the expected or appropriate response. Finally, the presence of others reduces the cost of non-intervention, as responsibility is diffused throughout the group. The knowledge that others are present allows the agent to shift some of the responsibility to them (Latané and Darley 1970, 125; Latané and Nida 1981, 309).

Contrary to the situationists' claim, however, this evidence does not show that the behavior of the agents in question did not stem from global dispositions. We should consider the possible role of vices such as selfishness, laziness, or cowardice. The three processes identified by Latané and Darley are perfectly compatible with the idea that such traits cause inaction under the circumstances. If one has a lazy or selfish disposition, for example, one does not help when one believes one could get away with inaction, and one can more easily get away with inaction when others are present than when one is alone. The fact that there are others who could also act makes it possible to shift some of the responsibility to others; it allows one to offer the excuse that the situation was ambiguous or confusing (after all, the others did not act, so perhaps there was something they knew that the agent did not); and it also adds the potential social costs of embarrassment if one's intervention turns out to be misguided. On the basis of some or all of these considerations, agents who are selfish, lazy, or cowardly may well refrain from action. The availability of an explanation in terms of non-malicious vices implies that the group effect as such is in principle compatible with the existence of global traits.

A subset of experiments tested the group effect under conditions where there appeared to be a threat to all, including the test subject, and at first sight it might seem as if an explanation of the sort we suggest is impossible in these cases. Latané and Darley (1968) had a room gradually fill with smoke. Ross (1971) and Ross and Braband (1973) set off a ringing
bell and a flashing 'Fire' sign. In these cases there was also a marked drop in intervention rates when subjects in the room were in a group compared to subjects who were on their own (Latané and Nida 1981, 312). Situationists might regard the results of these studies as incompatible with an explanation in terms of non-malicious vices. After all, one would be inclined to associate vices such as selfishness and cowardice with self-preserving behavior, so one might expect subjects with such vices to be more proactive in responding to the threat. Yet insofar as the inaction of others causes ambiguity about the situation, this may make the situation seem more ambiguous and less of a threat, which could explain the reduced intervention rates. And insofar as the presence of others diffuses responsibility, this may enable selfish or cowardly subjects to remain conservative with their efforts and not risk unwarranted intervention. In sum, the group effect is compatible with – and indeed perhaps best understood in the light of -- a range of global non-malicious vices, even when the scenario involves a threat to the test subject.

3.1.3. The Mood Effect

Situationists also ground their argument in experimental findings that indicate that people’s willingness to help is affected by changes in their mood. Interestingly, it turns out that being in a good mood or a bad mood (as compared to a neutral one) makes one more likely to help (Schaller and Cialdini 1990). It has also been demonstrated that mood and helping behavior are highly susceptible to minor situational influences, such as smell (Baron and Thomley 1994; Baron 1997), noise (Mathews and Canon 1975), minor good-fortune (Isen and Levin 1972; Levin and Isen 1975), and the weather (Cunningham 1979).

Situationists take these studies to indicate that people typically do not possess global character traits. After all, the factors that influence helping behavior are not themselves morally relevant, so the variation cannot be explained by reference to virtuous dispositions on
the part of the subjects, and malice does not seem to be operative either. This then leads them to regard this evidence as providing support for the thesis of situationism (e.g., Merritt et al. 2010, 356-357).

Again, however, this conclusion does not follow. Remarkably, the leading explanation of why negative moods lead to increased helping behavior – the ‘mood management hypothesis’ – and one of the two leading models for the effect of good mood on helping\(^5\) – the ‘mood maintenance hypothesis’ – both see increased helping as a way of benefiting the agent (Miller 2009, 152, 159). Helping behavior is associated with the rewards of praise and social status, which increase positive affect. When we feel bad we form a motivation to improve our mood, and helping is one way to do this. When we feel good, we form a motivation to maintain our positive affect, and again, one way to do this is by helping other people. Whether or not we conceive of this as a conscious process, the mood maintenance hypothesis makes clear that the behavioral evidence as such does not force us in the direction of situationism. For it provides an explanation in terms of an underlying disposition to help others only when doing so will make oneself feel better.

This analysis is compatible with apparent counter examples. Mathews and Canon (1975, 574-575) found that people are much less likely to help an apparently injured man pick up his books when there was a power lawnmower running nearby than when background noise levels were normal. At first glance this may seem to tell against the idea that people help in order to improve a negative mood (in this case due to the loud and unpleasant noise). Yet the mood-management hypothesis suggests that when helping is an effective means to improving mood, and when there is no less costly means available to do so, helping behavior will increase. In this experiment there was an easier means - escaping the noise by moving away. Given that the negative mood was a product of the loud noise and that it was easy to escape the noise, helping was not the easiest means for an agent to improve affect. We do not
claim that the mood maintenance hypothesis indeed provides the best explanation; there are other possibilities as well. Our point is simply that the behavior observed in these experiments permits an explanation in terms of a stable desire for positive affect, which means that the evidence as such does not clearly show that people are acting at the mercy of situational stimuli.

In sum, the results of the three groups of studies discussed in Section 3.1, which situationists regard as strong evidence in support of their position, can also be explained in a way that is consistent with the widespread possession of global character traits. Therefore, these empirical studies do not make the case against global traits in general. In other words, they do not make the case for situationism.

3.2. Moral Hypocrisy

There is at least one important line of research that lends initial empirical plausibility to the idea that people regularly act in ways that are consistent with non-malicious vices. Daniel Batson and colleagues have run a large number of empirical studies that show that many people display 'moral hypocrisy' (for representative work see Batson et al. 1997; Batson et al. 1999; Batson et al. 2002; Batson et al. 2006), that is, that people wish to appear moral without being willing to do what is morally required. Of the many studies in support of this explanation, we discuss only one here. Batson et al. (1997) asked subjects to assign two tasks – one with positive consequences and one with neutral consequences – to themselves and another participant in the experiment. Subjects were told that the other participant would believe that the assignment was random. Their instructions included the following text, which was intended to give them an explicit cue about the moral nature of the dilemma:
Most participants feel that giving both people an equal chance – by, for example, flipping a coin – is the fairest way to assign themselves and the other participant to the tasks (we have provided a coin for you to flip if you wish). But the decision is entirely up to you. (Batson et al. 1997, 1341)

Subjects were then given a coin to use in such a procedure if they wished, and they were left alone in a room. Afterwards, nearly all the participants said that assigning the positive task to the other person, or flipping the coin to decide, was the morally right thing to do, yet only about half chose to flip the coin. Of those who did not flip, 80-90% chose to assign the positive task to themselves. Even more interesting, however, is the finding that of those who did flip the coin 85-90% assigned themselves the positive task (Batson et al. 1997, 1342). In further studies of this nature (see for example Batson et al. 1999; 2002), too, participants who flipped the coin in private assigned the positive task to themselves in the vast majority of cases.

These studies are not by themselves sufficient to prove that most of the test subjects were in fact acting from global vices such as dishonesty and selfishness, as this would require testing specifically for cross-situational consistency in individual agents. But the evidence is very suggestive and clearly compatible with the widespread existence of such global vices.

3.3. Is There Direct Evidence Against Global Vices?

We have been arguing that the situationists have overlooked the explanatory value of non-malicious vices because their conception of vices is too narrow and limited to malice. There are a few experiments, however, that could be taken as providing counter-evidence against the suggestion that global vices are widespread. In these experiments, subjects often act admirably. Situationists take this evidence, in combination with studies in which subjects largely act deplorably (such as the Milgram experiments), to constitute evidence against the
existence of global character traits. For if some situational factors prompt people to act deplorably, and other factors prompt them to behave admirably, they argue, this constitutes evidence not just against agents possessing global virtues, but also against agents possessing global vices. Therefore, we should consider those experiments that situationists regard as supporting their view and that show agents displaying behavior that is generally regarded as morally good. A first objection to our thesis, then, is that some behavior observed in psychological experiments is simply too good for the agents to have global vices.

Peter Vranas cites two experiments in support of the claim that 'there are many situations in each of which most people (would) behave admirably' (Vranas 2005, 4). In the first (Clark and Word 1974), nearly all subjects helped a technician who had apparently suffered a severe electric shock. 'Helping' here meant either offering direct personal assistance or indirect assistance by reporting the incident or obtaining help from others. In the second (Moriarty 1975), subjects who had been asked by a confederate to watch some item (a bag, for example) were very likely to directly confront a confederate thief who took the item. Vranas argues that people’s willingness to risk harm from confronting the thief or helping the technician represents 'admirable' behavior. But if people behave, or would behave, admirably in many situations, he argues, we cannot give their overall character a negative evaluation: they are not ‘bad people’. On his account, people are neither good nor bad; they are indeterminate and have 'no character status' at all (Vranas 2005, 2).

Do cases like ‘thief and ‘technician’ indeed show that people's behavior is too good for them to possess global vices of that sort we have been discussing? We do not think so. First, we have not argued that people have only vices or that they only ever act on vices. There may well be a significant number of agents who do not have vices that impair their helping behavior in the two experiments at issue. If they have virtues, and even if their
behavior in these specific experiments is impossible to explain in terms of global traits at all, this is still compatible with global vices being rather common.

Second, cases like 'thief' and 'technician' would tell against the possession of non-malicious global vices only if the subjects' behavior revealed a motivation that was inconsistent with the possession of such vices. Even subjects with such vices will be led to the morally right thing under certain circumstances, however, and for all we know, this may have been true of a non-negligible number of the test subjects in these cases. In the case of ‘thief’, the subjects are directly asked to perform a task – watching and if necessary protecting the possessions of a confederate – and they explicitly agree to do so. They then find themselves faced with precisely the task they agreed to perform. Failure to do as they promised risks a confrontation or a serious loss of social standing if the confederate returns to find their possessions stolen and publicly blames the subject for it. Being motivated to avoid this kind of repercussion could explain—in at least a non-trivial number of cases—why people confront the thief. A similar kind of explanation might be available for ‘technician’. In this case, the subjects risk serious consequences if they walk past an obvious and life-threatening emergency without even reporting it, and this may motivate even a selfish person to do something. A lazy person may spring into action when the need is pressing enough, and reporting the incident is not so much work. A coward may enlist others to assist in helping the technician, but he will try to avoid touching the electrical equipment. These are just some examples of how the observed behavior could be explained in terms of underlying global traits. We are not claiming here that the helpful subjects in these studies were actually motivated by selfishness, laziness, or cowardice. The available evidence is insufficient to establish the motivations of the individual test subjects with certainty. As long as such non-malicious vices remain a possible (and prima facie not at all implausible) explanation of the subjects' behavior, however, the observed behavior does not count in favor of situationism.
A second objection against our thesis, somewhat different from the first, is that the behavior observed in test subjects is not bad enough to be explained in terms of vices. Christian Miller points to evidence of widespread cheating behavior (see Miller 2014, 62-65, for an overview), and widespread lying behavior (Miller 2013, 286-290), but he argues that this is not evidence that people possess a global trait of dishonesty because there are limits to the kinds of cheating and lying people engage in. In particular, while most people tend to cheat to some extent when the opportunity arises, they do not tend to cheat as much as possible. For example, Lisa Shu and colleagues (Shu et al. 2011) found that subjects answering problems, who would receive $0.50 for each correct answer, reported that they answered 13.22 problems correctly (on average) when they could shred the answer sheet immediately afterwards, so no one could check their result. This compared to an average of 7.97 correct answers in the control group. So while people clearly took the opportunity to cheat, they did not maximise their profit, for they could have reported a total of 20 correct answers. In further work, Nina Mazar and colleagues (Mazar et al. 2008) found that when subjects were prompted to recall the Ten Commandments, this kind of cheating went down. Miller follows Mazar and colleagues in holding that people typically have a conception of themselves as honest, and that this limits their cheating behavior (Miller 2013, 69). Moreover, he holds that the fact that people cheat less when primed by some moral code is evidence that people have the belief that honesty is appropriate, which can rule their behavior when they are primed in this way. Overall, Miller argues that there are a number of ways in which people act, appear to be motivated, and think of themselves that do not fit the pattern we would expect from people who have a global trait of dishonesty. In particular, we would not expect dishonest persons to think of themselves as honest, to hold that being honest is appropriate, or to fail to maximise a cheating opportunity (see Miller 2013, 312-313, for a summary of these conditions).
Miller's conception of vice is dissatisfying in several respects, however. First, maximising every cheating opportunity is not a plausible condition for the trait of dishonesty. Whether one lies that one had 13 or 20 correct answers, when one actually had only 8, one is cheating either way. Those who cheated may well have figured that it would be rather improbable, in the eyes of the experimenter, for them to have answered all questions correctly, and they may have wanted to avoid looking suspicious. Similarly, one can be disposed to lie without this requiring that everything one ever says be a lie; the disposition is usually tied to situations in which lying is to one's advantage and one can lie without ruining one's reputation as trustworthy or running into other trouble. The amount of lying a liar can engage in is naturally limited by the condition that one needs to preserve the trust of others for one's lies to be effective.

A second problem concerns Miller's contention that an agent, in order to qualify as having a vice, needs to be wholeheartedly committed to the behavior in question and not regard it as morally wrong (Miller 2013, 303, 312-313). This too seems an implausible condition. It is not odd to conceive of dishonest people as regarding themselves as being committed to honesty; after all, they are dishonest, so they may well engage in some form of rationalization. More important, however, is the fact that dishonest people need not endorse their own dishonest behavior in all respects in order to qualify as dishonest. They may harbor hopes of becoming better people--indeed they may feel bad about themselves--and this is compatible with their nevertheless robustly acting dishonestly across a broad range of situations. Indeed, within the tradition of Aristotelian virtue ethics as well as Kantianism, vices are usually regarded as dispositions that involve internal conflict on the part of the agent who has them (Aristotle 2004, 1166b; Kant 1996, 4:454). Miller rejects this aspect of vice, because he regards it as a necessary condition for attributing a vice to an agent that the agent feels no distress when acting in accordance with the vice and does not believe that
acting in this way is wrong. We regard this condition as too strong. A selfish agent may reliably decide to act in a way that is morally impermissible while knowing that it is morally wrong. For example, she may reliably decide to give priority to trivial interests of her own over the urgent needs of others, while knowing that her practical decisions are morally indefensible. Because such a 'selfish' agent is disposed to actively pursue her self-interest in morally problematic ways, her action is not due to weakness of will but to vice. The fact that she is aware that what she is doing is morally wrong does not make her action any less vice-based.

At the end of this discussion, it is worth noting that one should expect an important evidential asymmetry between virtues and vices. Virtues are, by their nature, praiseworthy, so we can expect people to express their virtues in public (at least insofar as the virtues they acknowledge are also regarded as such in their social and cultural context). Vices, by contrast, are traits people will often try to hide from the public eye. Therefore, we can allow more numerous instances of vice-contrary behavior before having to conclude that a person does not have vices than we can allow instances of virtue-contrary behavior before having to deny that a person has virtues.

This does not immunise vice from empirical challenge, however. If Batson's test subjects dutifully flipped their coins and assigned the positive consequences tasks to others in roughly 50% of the cases, we would have no reason to suspect them of moral hypocrisy. Rather, our point is that the evidential asymmetry between virtue-based and vice-based behavior should be taken into account when assessing the empirical evidence concerning the existence of global vices. It is much harder to rule out the widespread possession of global vices than situationists tend to assume.

Finally, we have not argued or meant to suggest that people's behavior is typically (let alone always) caused by global non-malicious vices. We have argued that situationists have
overlooked the possibility that the observed behavior is often (or at least in a substantial number of cases) best explained by reference to such vices and hence that their own position lacks the empirical support they claim for it. To what extent global non-malicious vices can actually explain human behavior remains to be seen, and establishing this requires research of a different type than the experiments we have been discussing. But even if longitudinal studies of large numbers of individual subjects showed that such vices explain behavior merely in a sizeable minority of cases, this would already suffice to confirm that situationism is mistaken. The situationists argue that global character trait-dependent behavior is 'rare enough to count as abnormal', and that people's morally relevant behavior is 'typically' the result of depersonalized response tendencies triggered by morally irrelevant or insignificant features in the situational context. To show this bold thesis to be mistaken, it is enough if global vices such as laziness, cowardice, and selfishness turn out to be common enough to count as all-too-ordinary elements of human moral psychology.

**Conclusion**

We have argued that situationists overlook the explanatory potential of global non-malicious vices, and that the empirical evidence to which they appeal is consistent with the widespread possession of global character traits because it is consistent with the widespread existence of global non-malicious vices. Our point is most fundamentally a point about the structure of the situationists' argument and their use of the empirical evidence. If our analysis is correct, the evidence does not show the thesis of situationism to be true; in fact, the evidence is fully compatible with the view they oppose, namely, the thesis that human behavior is often best explained by reference to global character traits.

Establishing the extent to which people in fact act on the basis of global traits--and *a fortiori* establishing the possibility of moral self-improvement and genuine virtue--is an
altogether different matter. Any descriptive claims regarding people's possession of virtues and vices will have to be based on a type of research that hardly exists and that is methodologically difficult to carry out, namely, on longitudinal studies involving the same subjects in a variety of circumstances.

References


**Notes**

---

1 There is something misleading about this second premise, because in a sense the situationists themselves argue that systematic observation does reveal pervasive behavioral consistency. They claim that human behavior reveals pervasive *situation-dependent* consistency. Therefore, it would have been clearer if they had formulated the second premise in terms of the absence of *trait-dependent* behavioral consistency. For, as Doris puts it, 'the
question is whether the behavioral regularity we observe is to be primarily explained by reference to robust dispositional structures or situational regularity (…). I insist that the striking variability of behavior with situational variation favors the latter hypothesis' (Doris 2002, 26). Thus, the question is not so much whether there is behavioral consistency, but whether the behavioral evidence is consistent with its being caused by underlying global traits. We shall take the second premise to mean that systematic observation does not reveal pervasive trait-dependent behavioral consistency.

2 Some situationists explicitly allow for the existence of such a minority, e.g., Vranas 2005, without, however, regarding this as evidence of global traits. On Vranas' view, it is compatible with the thesis that character is typically fragmented.

3 The point about the situationists' argument structure is developed in more detail in Kleingeld 2015, along with a critique of their conception of 'situation management'. The present essay shows how the results of the relevant empirical studies can indeed be interpreted as consistent with the widespread possession of non-malicious vices. A more extensive discussion of the empirical record can be found in Bates 2016.

4 It is worth noting that these data are compatible with, though not decisive evidence for, the existence of a disposition to cruelty in a small minority.

5 The alternative model is the 'concomitance model', which suggests that increased helping is a by-product of the other cognitive changes caused by positive affect, such as improved task performance, cooperativeness, optimism about the future, and increased information acquisition (Miller 2009, 154).
This is not to deny that there may be certain social contexts in which it is considered 'cool' to flaunt one's vices. Also, there may be significant disagreement as to whether certain traits are virtues or vices.