

SITUATIONS AND RESPONSIVENESS TO REASONS*

Carolina Sartorio

University of Arizona

Some classical studies in social psychology suggest that we are more sensitive to situational factors, and less responsive to reasons, than we normally recognize we are. In recent years, moral responsibility theorists have examined the question whether those studies represent a serious threat to our moral responsibility. A common response to the “situationist threat” has been to defend the reasons-responsiveness of ordinary human agents by appeal to a theory of reasons-responsiveness that appeals to patterns of counterfactual scenarios or possible worlds. In this paper I identify a problem with that response and I offer a better solution.

1. A Case

In a famous experiment in social psychology by Darley and Batson (1973), the “Good Samaritan” experiment, seminary students were asked to give a talk across campus. On the way there, they ran into a man who appeared to be (but wasn’t really) in need of urgent medical care. The study found that whether the students were likely to help heavily depended on how much in a hurry they were: if they had been told that they had very little time to get to their lecture (the “rush” condition), they were much less likely to help than if they had been told that they had more time (the “no-rush” condition). In recent years, moral responsibility theorists have focused on studies of this kind to see what implications they may have about the responsibility of human agents. In particular, they have examined whether the behavior of agents in studies of this kind suggests that we are less responsive to reasons, and thus arguably less morally responsible, than we normally take ourselves to be (see, e. g.,

Nelkin 2005, Nahmias 2007, Brink 2013, Vargas 2013, Stammers 2016, McKenna and Warmke Forthcoming, and Herdova and Kearns Forthcoming).

Conditions such as the rush and no-rush conditions in the Good Samaritan experiment are known as *situational* features. Situational features are external factors of the environment that don't make a difference to what we should do in the circumstances, but that still tend to have an effect on what we actually do. If you run into a man who desperately needs your help, you should assist him, even if that means that you'll be late for a lecture. (Although being in a rush could provide some reason not to help, it is not a sufficient reason not to help, in the circumstances, and it doesn't make a difference to what you should do in that case.) In some cases, the studies also seem to show that situational factors can affect the behavior of human agents without our even realizing that this is the case; in fact, by contradicting the principles of action that we consciously stand by (such as the principle that we should assist someone in need even if this means that we'll be late for a lecture). Thus the threat to our responsibility posed by studies of this kind is that they seem to show that we are more sensitive to situational factors, and less responsive to reasons, than we normally recognize we are. I'll refer to this challenge as *the situationist threat*.

Several reasons-responsiveness theorists have pushed back against this threat. They have pointed out that studies like the Good Samaritan experiment fail to show that agents are prey to situational factors, at least in any radical kind of way. For, they have argued, even in those cases agents tend to retain a sufficient extent of responsiveness to reasons (see Brink 2013, Vargas 2013, McKenna and Warmke

Forthcoming, and Herdova and Kearns Forthcoming). In general terms, the strategy has been the following. According to traditional theories of reasons-responsiveness, whether agents retain a sufficient extent of responsiveness to reasons is a matter of whether they had on the relevant occasion the capacity to respond to the relevant reasons, regardless of whether they actually exercised such a capacity. And, even though some agents who are subject to situational influences don't respond to the actual reasons due to the presence of those factors, they retain the relevant capacity to respond to those reasons.

What determines whether those agents have or lack the relevant capacity? On these views, the answer is given by the existence or nonexistence of certain *counterfactual* scenarios, or of a pattern of counterfactual scenarios, where the agents respond to the relevant reasons and thus act otherwise. If there are enough possible worlds, or patterns of possible worlds, where the agents respond to the relevant reasons, this is sufficient for those agents to have the capacity in question, and thus to be responsible. How many of those counterfactual scenarios are "enough" is, of course, not something that can be given a precise characterization, and it may depend on the nature of each specific case. But the thought is that for each set of circumstances we'll be able to identify some relevant set of counterfactual scenarios and some vaguely specified threshold beyond which we would judge that the agent had the capacity in question.

So, to return to the Good Samaritan case, the question would be whether the agents who failed to help the victim in the rush condition still retained the capacity to help him and thus can be morally responsible for not helping him. And this would

be a matter of whether there are enough counterfactual scenarios, or patterns of counterfactual scenarios, where those agents respond to the relevant reasons and act otherwise.

What could be such counterfactual scenarios? Well, for one thing, notice that the experimental results themselves suggest that many of the subjects who fail to assist the man in the rush condition would have in fact assisted him if they had been placed in the no-rush condition instead. After all, what the study is interpreted to suggest is precisely that whether or not you're in a hurry is likely to make a difference to what you do in those circumstances. So this means that there are at least some counterfactual scenarios, scenarios where the agents are not in a hurry, where the agents in question tend to respond to the relevant reasons (Herdova and Kearns Forthcoming).¹ Also, the experimental results are compatible with the existence of several other scenarios in which the agent is equally in a rush but stops to help the victim. For example, these could be scenarios in which the victim was much more visibly hurt, such as someone who had an axe sunk in their skull (McKenna and Warmke Forthcoming). It is reasonable to expect that many of the original subjects in the experiment would have helped the victim if they had been placed in such an "axe-in-skull" condition, even if they were in a hurry. And the same can be said about several other non-actual conditions. (Of course, we could only know for sure by running more studies! But note that, in order to test for this, we would have to run lots of studies in which we'd vary different aspects of the circumstances that we think are likely to have an influence on the agents' behavior.)

So, if there is such a variety of counterfactual scenarios or pattern of possible worlds where the subjects in studies like the Good Samaritan case respond to the relevant reasons and assist the victim, it follows from traditional theories of reasons-responsiveness that those agents are sufficiently responsive to reasons and thus morally responsible for actually failing to assist him. Of course, they could still be somewhat *less* responsive to reasons than other subjects who were never placed in those conditions. And, as a result, assuming that moral responsibility comes in degrees and that it tracks degrees of reasons-responsiveness, they could also be somewhat less responsible than those other subjects. But the main point is that they would still be morally responsible for their behavior. The situationist threat has been defused, for the most part, or so it seems.

2. A Problem

But there is a basic worry with this strategy. The worry can be motivated in the following way. The situationist threat arises from the fact that we can apparently be “gripped” by some aspects of the actual circumstances (in many cases, without our even realizing this) in ways that can undermine our control and responsibility. Thus, in order to determine whether this is true in a certain case, the objection goes, we cannot suppose away those very aspects of the circumstances and try to figure out how the agent would have behaved in scenarios where the circumstances were different. For, if we’re trying to determine whether the circumstances—the *actual* circumstances—are gripping us, it’s of no use to look at what we would have done if those circumstances had been different.

Compare: there are other things that can (much less controversially) grip us in a way that they can rob us of control and responsibility. A compulsive desire, for example, can have such an effect on us. Imagine an extreme addict who is irresistibly compelled to take a drug on a certain occasion. An account of responsibility in terms of reasons-responsiveness would say that the addict is not responsive to reasons when he acts on his desire to take the drug on that occasion. Now, imagine arguing that this is not actually true because the addict retains the capacity to fail to take the drug *when he doesn't desire to take it*—say, in circumstances where he took the drug very recently. Of course, in most of those circumstances where he doesn't desire to take the drug, he doesn't take it. Does this mean that our addict retains some measure of reasons-responsiveness?

Of course not! Circumstances where he doesn't desire to take the drug just aren't relevant to how reasons-responsive he is when he takes the drug driven by his desire to take it—in other words, when his addiction kicks in. If we are trying to figure out whether a behavior by an agent is a compulsive behavior, what we are trying to determine is whether his actual desire grips him in a way that robs him of control and responsibility. In order to answer this question, we cannot look at cases where the desire isn't present, but must focus only on what happens to the agent when the desire is present. After all, the compulsion isn't always at play, but is triggered by the desire (or by the circumstances giving rise to the desire), and what we are interested in finding out is whether the agent acts freely in those specific circumstances, where he is driven by the desire.²

The situationist threat must be interpreted along similar lines, I claim, lest we underappreciate its force. The contention is that external circumstances can grip us too, in a similar way, and can also rob us of control and responsibility. But then it seems that we cannot assume that the agent's reasons-responsiveness is determined by what she does in conditions where those very same circumstances are absent. In particular, blaming the agents in the rush condition for not helping the victim on the basis of what they would have done if they had been placed in the *no-rush* condition seems wrong, just as wrong as blaming the addict for succumbing to his desire to take the drug on the basis of the fact that he can easily refrain from taking the drug whenever he doesn't desire to take it. After all, what we are trying to figure out is whether those agents in the rush condition were sufficiently gripped by the specific circumstances that are the *actual* circumstances, in particular, by their being in a hurry. What they do in scenarios where the circumstances are different—circumstances where they are not in a hurry—is arguably irrelevant to this question.

In response, it could be argued that reasons-responsiveness is a matter of degree, and that, even in the case of the addict, he is *slightly more* sensitive to reasons given what he does in possible worlds where he doesn't desire to take the drug. So those worlds do count towards his reasons-responsiveness, although they are probably not enough to make him responsible. On this view, *all* worlds (or, perhaps, all nearby worlds) where agents have sufficient reasons to refrain from acting and where they act on those reasons count toward their reasons-responsiveness.

But this is to miss the point of the objection. The objection is that, in the case of the addict, possible worlds where he doesn't desire to take the drug are just plain *irrelevant* to his reasons-responsiveness on that occasion (similarly, in the Good Samaritan case, possible worlds where the agents aren't in a hurry, are just plain irrelevant to their reasons-responsiveness on that occasion). They are irrelevant because they don't speak to how responsive to reasons he was when he took the drug *while being under the grip of the desire*, which is what we want to know in that case. So those possible worlds shouldn't even be among the relevant possible worlds, when we are assessing the reasons-responsiveness of the agent in those cases. Taking reasons-responsiveness to be a matter of degree is consistent with thinking that there are possible worlds where the agent has sufficient reasons to refrain from acting in the way he does in the actual world, but that are simply irrelevant to his reasons-responsiveness on that particular occasion.

Imagine, then, that we are moved by the force of this argument and thus we concede that possible worlds where the agents are not in a hurry are just irrelevant to their reasons-responsiveness in the rush condition. How about other possible worlds, such as (to use the same example by McKenna and Warmke from before) worlds where the victim lies on the ground with an axe sunk in their skull? Couldn't those counterfactual scenarios be relevant?

Yes, in principle, they could. But reasons-responsiveness views face a *demarcation problem* at this point. For we need some principled reason to single out the aspects of the actual circumstances that we can vary from the aspects of the circumstances that we must hold fixed in order to assess an agent's reasons-

responsiveness on a certain occasion. Once we acknowledge that *not all* possible worlds where the agents have sufficient reasons to refrain from acting are relevant to their reasons-responsiveness in the actual scenario, we need to say more about which ones are relevant and which ones aren't. In the Good Samaritan case, for example, what determines whether the agents in the rush condition were morally responsible for failing to assist the victim? Is it their reasons-responsiveness just holding fixed that they were in a hurry? Or is it their reasons-responsiveness *also* holding fixed, for instance, that there was no extremely vivid emergency such as the presence of a victim with an axe sunk in her skull?

These are hard questions. But unless we can figure out how to answer them we won't be able to settle the issue of whether the agents were sufficiently responsive to reasons on that occasion. For the extent of their responsiveness to reasons will vary significantly, depending on which aspects of the circumstances we hold fixed and which ones we vary. So there is an interesting puzzle that arises for reasons-responsiveness views, one that the situationist threat helps bring out in a particularly vivid way. A bit more precisely, the puzzle is this. Examples where situational features have an influence on the behavior of agents can be used to illustrate the fact that, in assessing an agent's responsiveness to reasons in counterfactual terms (in terms of their behavior in counterfactual scenarios), at least *some* aspects of the actual circumstances must be held fixed. For, if we don't hold any aspects of the circumstances fixed, we face the irrelevance worry: the worry that the counterfactual scenarios we are focusing on are just irrelevant to the agent's responsiveness to reasons in the actual circumstances. Now, of course we

can't hold *all* the aspects of the circumstances fixed, on pain of trivializing reasons-responsiveness views. So we need some principled criterion to set apart, in each case, the aspects of the circumstances that we should hold fixed from those that we can legitimately vary. Unless we can offer such a criterion, reasons-responsiveness views rest on shaky foundations.

This is an important problem for traditional theories of reasons-responsiveness, one that deserves immediate attention. Surprisingly, very little has been said in the literature about it; at least, very little has been said that speaks directly to this issue.³

I will argue for a radical solution to the problem, one that is so radical that it manages to avoid it altogether. I will suggest that we can avoid the demarcation problem by switching to an *actualist* conception of reasons-responsiveness. It is an "actualist" conception in that it understands responsiveness to reasons, not in terms of patterns of counterfactual scenarios, but just in terms of the actual scenario and the actual circumstances in which the agent acts. Of course, if we just focus on the actual scenario, instead of on departures from the actual circumstances or in merely possible worlds, the problem of demarcation simply disappears. For then we don't have to worry about which variations are allowed and which aren't, or which possible worlds are relevant and which aren't.

3. A Solution

But how, you may ask, could responsiveness to reasons *just* be a matter of what happens in the actual scenario? At first this may sound completely confused, almost like a contradiction in terms.

But it's not. The basic idea is this.⁴ In cases where agents are responsive to reasons, there are many aspects of the actual circumstances that are explanatorily relevant to the agents' behavior but that tend to remain unnoticed. In the Good Samaritan case, one of them has to do with the axe-in-skull condition discussed above. Imagine that, as it is reasonable to expect, most subjects in the rush condition are "ordinary" human beings in that they are sensitive to extremely vivid emergencies such as an axe's sticking out from the victim's skull. Call one of those agents "Sensitive." Now imagine that there is also an (out-of-the-ordinary) agent in the group, call him "Insensitive," who is so gripped by the rush condition that he becomes insensitive to any such emergencies whenever he is in a hurry. Insensitive is, in an important sense, like our earlier example of the addict, who is gripped by his compulsive desire to take the drug and is insensitive to pretty much anything else that might be going on around him whenever he is under the influence of the desire: he stands to the circumstance of being in a hurry in the same way the addict stands to his desire to take the drug.

Again, imagine that both of our agents, Sensitive and Insensitive, fail to assist the victim when they are placed in the rush condition. Still, surely, there is an important difference between them. The difference I have in mind doesn't just concern how they behave in other possible worlds, but it concerns, more fundamentally, the *actual* explanation of their behavior. Plainly put, the difference is

that, in choosing not to assist the victim in the actual circumstances, Sensitive was responding to the absence of vivid emergencies and Insensitive was not.

Compare this other case. Imagine that you and I like to sing at the top of our lungs while showering, but you are much more considerate than me: you refrain from doing so whenever there is somebody in the house who doesn't appreciate your musical performance, but I don't care at all about that and do so regardless, without exception. Imagine that on a certain occasion where nobody is around, we both sing "Bohemian Rhapsody" at the top of our lungs. Surely, even in that case, there is a difference in the actual explanation of our behavior: you do it, among other things, because nobody is around, but I don't. I'm just insensitive to that aspect of the actual circumstances. So the fact that nobody is around is not part of the explanation of my singing; however, it is in your case, as in the case of any other equally considerate person.

This difference between the considerate and the inconsiderate singer is a real and significant one. But it may be easy to overlook it, at first glance, in scenarios of this kind, where the actual behavior is exactly the same, and where there is a more obvious or salient cause that partially explains the behavior in both cases (our desire to sing while showering). In particular, note that the difference concerns the role of an *absence* (the absence of an annoyed audience) in the explanation of our behavior. So it might be easy to overlook this difference, in these cases, because it has to do with the role played by an absence instead of a positive event. Absences tend to remain unnoticed, and for good reason: it's hard to keep track of all the absences in the world! Still, we would be missing something important about those

agents and their behavior if we were to say that they do what they do for exactly the same reasons. They *don't* do what they do for exactly the same reasons, since they are not being sensitive to all the same aspects of the circumstances.

You might be wondering how this difference between the explanation of the considerate singer's behavior and the explanation of the inconsiderate singer's behavior should be cashed out in metaphysical terms. This is a difficult issue, one that concerns how we should make sense of talk about absence causation or about the explanatory power of absences, and this is not the place to discuss it. Still, regardless of how exactly the difference should be cashed out, it seems clear that it is a real and important difference, as the examples of the considerate and inconsiderate singer help illustrate.⁵

I submit that we should analyze the behavior of our agents in the Good Samaritan scenario, Sensitive and Insensitive, in a similar way. The actual explanation of Sensitive's behavior includes the no-axe-in-skull aspect of the circumstances, or the absence of an axe in the victim's skull, whereas this isn't true of Insensitive's behavior, given that he is insensitive to those things (when he is in a hurry). Although the two agents in fact behave in the same way, and although some of the causes of what they do are the same (they both fail to assist the victim because they are in a hurry), they don't do what they do for exactly the same reasons, or for *all* the same reasons. Again, we would be missing something important about their behavior if we were to say that they behave in the way they do for exactly the same reasons, or for all the same reasons.

This generalizes to other aspects of the circumstances that could plausibly be playing a role in the agents' behavior. Whenever there is a difference in the agents' sensitivity to certain aspects of the situation, that difference will be reflected in the actual explanation of their behavior. This means that we don't need to look at merely possible worlds or counterfactual scenarios to locate a difference in the reasons-responsiveness of these agents: the actual world and the actual circumstances are sufficient, once we look close enough.

What we need to do, then, is build a theory of reasons-responsiveness that appeals to such features of the actual explanation of behavior. This is an actualist theory of reasons-responsiveness. Roughly, it is a theory according to which, whenever the actual explanation of the agent's behavior exhibits sensitivity to enough of the relevant aspects of the circumstances, the agent is reasons-responsive; otherwise he is not. For example, the view would say that Sensitive is sufficiently responsive to reasons and Insensitive is not because the actual explanation of Sensitive's behavior reveals that he is responsive to certain relevant aspects of the circumstances (the no-axe-in-skull condition, among many others), while the actual explanation of Insensitive's behavior does not.⁶

This view completely avoids the problem of demarcation that arises for traditional (i.e. counterfactualist) reasons-responsiveness views. Again, since it doesn't appeal to possible worlds or hypothetical changes in the actual conditions, there is no need to distinguish those possible worlds that are relevant to the agent's reasons-responsiveness from those that are not, or those aspects of the circumstances that we have to hold fixed in assessing the agent's reasons-

responsiveness from those that we can legitimately vary. All we should look at, according to the actualist view, is the actual explanation of the agent's behavior. In some cases, certain aspects of the actual circumstances will be explanatorily relevant to the agent's behavior, and in other cases they won't be. That's all there is to it.

The way I have motivated the actualist view can also be used to shed light on some remarks I made earlier, when I introduced the problem of demarcation by appeal to the Good Samaritan case. I noticed then that there seems to be something inappropriate about appealing to what the agents would do in the *no-rush* condition to assess their reasons-responsiveness in the actual, rush condition. Now we can see why. In the Good Samaritan scenario the no-rush condition *isn't actual*. As a result, on the actualist view, it just isn't relevant to the agent's reasons-responsiveness in the actual, rush conditions. Again, this strikes me as exactly what we want to say in this case. Just like the addict's reasons-responsiveness when he takes the drug driven by the desire to take it cannot be determined by what he does in scenarios where he doesn't have the desire, the reasons-responsiveness of an agent in the rush condition cannot be determined by what the agent does in the no-rush condition.

In contrast, notice that the no-axe-in-skull condition (that is to say, the *absence* of an axe in the victim's skull) *is* actual: as a matter of fact, there happens to be no axe sunk in the victim's skull, in the circumstances where the agents act. Not only is it actual, but, as we have seen, it can play an explanatory role in the agents' behavior (as it does in the case of Sensitive: the absence of an axe in the victim's

skull is part of the explanation of why Sensitive fails to assist the victim). As a result, it can be relevant, and arguably it is relevant, to the agents' responsiveness to reasons in the actual circumstances. Again, this strikes me as the right result. As illustrated by the contrast between Sensitive and Insensitive, whether an agent was being sensitive to conditions of that kind seems to matter to whether we would judge that she was responsive to reasons when she acted. When Sensitive acted, she seemed to be responsive to reasons in a way that Insensitive was not, given that she was responding to certain relevant aspects of the circumstances that Insensitive wasn't responding to.

So, on this view, many different aspects of the circumstances can be relevant to the agent's reasons-responsiveness—basically, all of those aspects of the circumstances that we would identify as reasons, or absences of reasons, that rationalize the agent's behavior. But they are all aspects of the *actual* circumstances. Although the appearance of an emergency is actual, the agents in the Good Samaritan case aren't responding to it, so this doesn't count in favor of their reasons-responsiveness. However, there are many other aspects of the actual circumstances that are intuitively also relevant to their reasons-responsiveness on that occasion, and that it *is* plausible to think they are responding to. It is easy to miss them, because they take the form of absences: the absence of an axe sunk in the victim's skull, the absence of yelling on the victim's part, etc. But, on reflection, they seem to be just as relevant to the agent's reasons-responsiveness as the positive ones. So, on this view, to the extent that agents are sensitive to these others aspects

of the circumstances when they act, they can still be sufficiently responsive to reasons.

At this point, traditional reasons-responsiveness theorists could think of appropriating these results and using them to answer the demarcation problem that arises for their views. They could claim, in particular, that possible worlds where the agents are not in a hurry are not relevant to their reasons-responsiveness, but possible worlds where the victim has an axe sunk in his skull (or where the victim is screaming loudly, etc.) are.

But notice what has happened here. This solution would simply piggyback on the motivation for the actualist view. And, once we have available the actualist account of these cases, there is *no further work* to be done by the possible worlds. Once we realize that there is an actual difference between our agents, Sensitive and Insensitive, a difference that concerns just the actual explanation of their behavior, we no longer need to appeal to possible worlds to account for the difference in reasons-responsiveness.

This is a point worth emphasizing. When I claimed that we would be missing something important about the agents' behavior if we were to say that Sensitive and Insensitive did what they did for exactly the same reasons, I meant that *everybody* should accept that there is a difference in the actual explanation of their behavior. In particular, traditional reasons-responsiveness theorists should accept this too. Moreover, everybody should accept that such a difference in the actual explanation of their behavior is intuitively relevant to their reasons-responsiveness. But then it seems that the appeal to counterfactual scenarios is no longer motivated, or needed.

To clarify this point further: this is not to say that counterfactual scenarios or possible worlds needn't play a role on the actualist view. They *could* still play a role, even on the actualist view, by grounding the claims about the actual explanations of behavior. For example, one could have the view that the fact that Sensitive is responsive to the absence of an axe in the victim's skull is grounded in the fact that he responds to the presence of an axe in some possible worlds where the victim has an axe sunk in his skull. Whether we should say that possible worlds play this kind of role would depend on what we think about how claims about actual causes or actual explanations are grounded (if we think that they are grounded in anything at all). If we believe that they are grounded in counterfactual facts, then the counterfactual facts become relevant by virtue of providing the grounds for the relevant facts about actual causes or actual explanations.⁷

I don't want to take a stand on this issue here, and I don't need to. For note that, even if some possible worlds were to play that kind of role, they would only do so insofar as they help ground the relevant facts about actual causes or actual explanations. The counterfactual facts would only be relevant *because of their connection with the actually explanatory facts*. So, on this view, possible worlds wouldn't play any role independently of this subsidiary role, and responsiveness to reasons would still just be a matter of facts about the actual sources of behavior (the facts about actual explanations, and whichever facts ground those facts).

4. Conclusion

The situationist threat is a threat to our reasons-responsiveness posed by external circumstances that are said to be so strong that they can deprive us of our reasons-responsiveness on certain occasions. When we are subject to such forces, the claim is, the situational factors “grip” us and render us blind or irresponsible to our reasons. In this paper I have outlined what I take to be the right answer to the situationist threat.

I have argued, first, that the threat cannot be answered simply by drawing on counterfactual scenarios that differ from the actual world with respect to the very aspects of the circumstances that are said to grip the agent. Instead, I have argued that we should answer the threat in a radically different way. Basically, the strategy consists in drawing on other, usually less immediately obvious, aspects of the *actual* circumstances, aspects that (it is reasonable to assume) the agents in question *are* responding to when they act in the actual scenario. I have suggested that this can be done by appeal to a plausible and independently motivated actualist theory of reasons-responsiveness.

The actualist view that I have sketched here has some important features that are worth mentioning. First, and most obviously, it comes apart from traditional (counterfactualist) views of reasons-responsiveness in that it disposes of the concept of responsiveness to reasons as a capacity that is possibly unexercised, and replaces it with a concept that is purely based on the actual sources of behavior. As I have explained, this has the advantage that it helps us avoid a serious problem for traditional views of reasons-responsiveness, which I have referred to as “the

problem of demarcation.” Once we embrace the actualist framework, the problem simply vanishes.

As an added bonus, the actualist view sits well with an increasingly popular and elegant conception of moral responsibility, the one spearheaded by Harry Frankfurt in his classical paper on responsibility and alternative possibilities (Frankfurt 1969). According to this conception of responsibility, the type of control required to be morally responsible for our behavior is just a matter of the actual sources of our behavior, or of why we actually do what we do; *nothing else matters*. The actualist framework of reasons-responsiveness that I have sketched here is a good match for this conception of responsibility.⁸

To conclude, then: what have we learned about the subjects in the Good Samaritan study, and about the reasons-responsiveness of human agents more generally? Are we really sensitive to the various aspects of the actual circumstances that we need to be sensitive to in order to be sufficiently responsive to reasons? At the end of the day, this is an empirical question, so the most I can offer here is an educated guess. But at least we can draw the following conclusion. Unless it can be shown that most of us are like our hypothetical character Insensitive, in that we are generally insensitive to the various aspects of the actual circumstances that are relevant to our reasons-responsiveness, we shouldn't conclude that studies like the Good Samaritan experiment prove that we are not responsive to reasons. Thus, unless and until that can be shown, we shouldn't regard the situationist threat as a serious challenge to our responsibility.

References

- Brink, D. 2013. "Situationism, Responsibility, and Fair Opportunity," *Social Philosophy and Policy* 30: 121-149.
- Darley, J. and C. Batson. 1973. "From Jerusalem to Jerico: A Study of Situational and Dispositional Variables in Helping Behavior," *Journal of Personality and Social Psychology* 27, 1: 100-108.
- Dowe, P. 2001. "A Counterfactual Theory of Prevention and 'Causation' by Omission." *Australasian Journal of Philosophy* 79: 216-26.
- Fischer, J. and M. Ravizza. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. New York: Cambridge University Press.
- Frankfurt, H. 1969. "Alternate Possibilities and Moral Responsibility." *Journal of Philosophy* 66, 23: 829-39.
- Haji, I. 1998. *Moral Appraisability: Puzzles, Proposals, and Perplexities*. New York: Oxford University Press.
- Herdova, M. and S. Kearns. Forthcoming. "This is a Tricky Situation: Situationism and Reasons-Responsiveness." *Journal of Ethics*.
- Lewis, D. 1986. "Causation." In *Philosophical Papers II*, New York: Oxford University Press, pp. 159-213.
- McKenna, M. 2005. "Reasons Reactivity and Incompatibilist Intuitions." *Philosophical Explorations* 8, 2: 131-43.
- McKenna, M. and B. Warmke. Forthcoming. "Does Situationism Threaten Free Will and Moral Responsibility." *Journal of Moral Philosophy*.

- Nahmias, E. 2007. "Autonomous Agency and Social Psychology." In M. Maraffa, M. De Caro, and F. Ferretti (eds.). *Cartographies of the Mind*. Berlin: Springer.
- Nelkin, D. 2005. "Freedom, Responsibility, and the Challenge of Situationism." *Midwest Studies in Philosophy* 29: 181-206.
- Sartorio, C. 2015. "Sensitivity to Reasons and Actual Sequences." In D. Shoemaker, ed. *Oxford Studies in Agency and Responsibility* (Oxford: Oxford University Press), volume 3, pp. 104-19.
- Sartorio, C. 2016. *Causation and Free Will*. Oxford: Oxford University Press.
- Stammers, S. 2016. "Situation, Reason, and the Extended Agent." *Teorema* 35, 2: 7-27.
- Vargas, M. 2013. "Situationism and Moral Responsibility: Free Will in Fragments." In J. Kirvenstein, A. Clark, and T. Vierkant, eds. *Decomposing the Will* (Oxford University Press), pp. 325-50.

* For helpful comments, thanks to Juan Comesaña, Michael McKenna, and an audience at Universidad Autónoma de Madrid.

¹ Herdova and Kearns use Fischer and Ravizza's account of reasons-responsiveness (Fischer and Ravizza 1998) as their main example of a reasons-responsiveness view. They apply Fischer and Ravizza's view to scenarios like the Good Samaritan and argue, on that basis, that agents remain sufficiently reasons-responsive in those cases. But they also point out that other accounts of reasons-responsiveness, such as Vargas' 2013 account, which relativize reasons-responsiveness to the agents' circumstances, would have different results in those cases. What I'll suggest in the next section is, in effect, that some relativization to circumstances is always warranted, given any plausible account of reasons-responsiveness.

² In this vein, Haji (1998) defends a theory of reasons-responsiveness according to which in order to determine how responsive to reasons agents are we need to hold fixed the agents' motivational state. Such an account would entail that the addict's compulsive desire has to be held fixed in order to assess his reasons-responsiveness.

³ McKenna 2005 is an exception. McKenna argues for a view according to which the best way to understand reasons-responsiveness (at least in the case of blameworthiness) is in terms of the capacity to respond to the *actual* moral reasons.

In order to do this, he argues, we must hold fixed those reasons while varying other aspects of the circumstances (all of this is within the context of Fischer and Ravizza's 1998 reasons-responsiveness account). It's an interesting question whether this suggestion by McKenna can be modified and suitably extended to the Good Samaritan case—where, as we have seen, one should arguably also hold fixed some salient features of the circumstances, such as the fact that the agent was in a rush. Instead of trying to settle this issue here, my strategy will be to argue for a different conception of reasons-responsiveness, one for which the problem doesn't even arise.

⁴ I develop this in more length in Sartorio 2016, chapter 4. Here I'm mostly interested in how the view helps avoid the problem of demarcation that arises for reasons-responsiveness accounts. See also Sartorio 2015, where I hint at a related issue (see especially section 5.5).

⁵ I discuss this issue and different possible approaches to it in Sartorio 2016, chapter 2. The main possibilities include: understanding the difference in terms of causation, causal explanation, or other relations like quasi-causation (Dowe 2001).

⁶ For more on this, see Sartorio 2016, chapter 4. Again, here I am not interested in giving the details of such a theory, but simply on its basic claims and motivations, and in how it helps with the demarcation problem faced by reasons-responsiveness views. The relevant aspects of the circumstances are reasons for action and *absences of reasons* for action. Note that, on this view, reasons-responsiveness still depends on a vaguely specified threshold: being responsive to reasons is a matter of actually responding to *enough of* the relevant aspects of the circumstances.

⁷ Here I have in mind counterfactual accounts of causation or explanation (see, e. g., Lewis 1986).

⁸ I argue for this in depth in Sartorio 2016. Of course, the actualist view could still work as an adequate theory of reasons-responsiveness even if the conception of responsibility motivated by Frankfurt failed. In other words, although the actualist view is a good match for such a conception of responsibility, it doesn't depend on the truth of such a conception of responsibility.