Routledge
Taylor & Francis Group

Inquiry
An Interdisciplinary Journal of Philosophy

ISSN: 0020-174X (Print) 1502-3923 (Online) Journal homepage: https://www.tandfonline.com/loi/sinq20

Situationism, subjunctive hypocrisy and standing to blame

Adam Piovarchy

To cite this article: Adam Piovarchy (2020): Situationism, subjunctive hypocrisy and standing to blame, Inquiry, DOI: 10.1080/0020174X.2020.1712233

To link to this article: https://doi.org/10.1080/0020174X.2020.1712233

INQUIRY
https://doi.org/10.1080/0020174X.2020.1712233

Routledge
Taylor & Francis Group

Situationism, subjunctive hypocrisy and standing to blame

Adam Piovarchy

The University of Sydney, Sydney, Australia

ABSTRACT

Philosophers have argued that subjects who act wrongly in the situationist psychology experiments are morally responsible for their actions. This paper argues that though the obedient subjects in Milgram's 'Obedience to Authority' experiments are blameworthy, since most of us would have acted in the same manner they did, it is inappropriate for most of us to blame them. On Todd's ([2019]. "A Unified Account of the Moral Standing to Blame." *Noûs* 53 (2): 347–374.) recent account of standing to blame, agents lack the standing to blame others for a wrong when they are not sufficiently committed to the moral values which would condemn that wrong. I argue that the obedient subjects lack sufficient commitment to the kinds of values which would condemn their wrongdoing. This is evidenced by the fact that the wrongdoing was severe, that the subjects had the capacity to avoid wrongdoing, and that there was very little cost to avoiding wrongdoing. Since these studies are very well-replicated, most of us in the moral community would have acted as they did for similar reasons. At least 80% of us therefore lack the standing to blame the obedient subjects.

ARTICLE HISTORY Received 12 August 2019; Accepted 25 November 2019

KEYWORDS blame; hypocrisy; situationism; standing; Milgram

1. Introduction

The situationist psychology experiments are a suite of studies demonstrating that normatively irrelevant features of a situation can have large impacts on people's moral behaviour. These experiments are rather unsettling, because they demonstrate that a large proportion of people can readily be made to act immorally in situations where doing the right thing appears easy. While most philosophical analysis of these experiments has focussed on whether these studies show that our common attributions of broad-based character traits (e.g. bravery or compassion)

CONTACT Adam Piovarchy adam.piovarchy@sydney.edu.au Office A14, Department of Philosophy, School of Philosophical and Historical Inquiry, Faculty of Arts and Social Sciences, The University of Sydney, 2006, NSW, Australia

are empirically indefensible (Doris 2002; Solomon 2005; Sreenivasan 2002), a number of philosophers have examined whether subjects who act immorally in these settings are morally responsible for their wrongdoing (Brink 2013; McKenna and Warmke 2017; Nelkin 2005; Talbert 2009; Vargas 2011). The number of subjects who act immorally in these settings conflicts with our folk predictions of how most people would act. This, along with observations of which kinds of factors influence subject behaviour, leads some people to think that the subjects who act immorally in these settings may be excused for their actions.

Though I believe that the subjects in these experiments are morally responsible for their actions, I also believe that the common intuition that they should be excused still tracks something important. This regards whether we have the standing to blame the subjects who act immorally. Specifically, I believe that though the subjects are blameworthy, it would nevertheless be inappropriate for us to blame them because we likely would have acted in exactly the same manner they did.

This paper proceeds as follows. First, I provide a brief summary of the situationist experiments and philosophical work examining whether subjects who act wrongly in these settings are morally responsible and therefore blameworthy for their actions. Second, using Todd's (2019) account of standing to blame, I argue that if an agent is not sufficiently committed to the values which would condemn a wrong, that agent lacks the standing to blame other agents who commit similar wrongs. This entails that if an agent would commit an easily avoidable wrong due to a lack of commitment to the relevant values, that agent lacks the standing to blame agents who in fact commit that wrong. I then argue subjects who act wrongly in Milgram's obedience to authority experiments should be thought of as having insufficient commitment to the relevant values, given the low cost of avoiding wrongdoing and their capacity to disobey. Since these experiments are generally well-replicated in many different populations, I conclude that approximately 80% of us lack the standing to blame subjects who obeyed in Milgram's experiments.

2. The situationist experiments

There are hundreds of situationist experiments demonstrating that environmental factors which seem to be morally and prudentially irrelevant have substantial effects on subject behaviour (see Doris 2002, 23–53, for a summary). Of particular interest has been a group of experiments demonstrating that the presence of other agents and perceived social pressure

can lead large numbers of people to commit wrongdoing. Darley and Latane (1968) demonstrated that having four people present, but not helping, reduces the chances an agent will help someone audibly having an epileptic seizure in the next room from 85% to 31%. Darley and Batson (1973) demonstrated that being told one is slightly late to give a talk reduces the chances individuals will check on someone who appears to be unconscious from 63% to 10%.

The most well-known situationist experiments, which I will be focussing on in this paper, are Milgram's (1963) obedience to authority experiments. In these experiments, subjects sign up to what they believe is an experiment on memory and learning. The experiment has two roles, teacher and learner. A machine which delivers electrical shocks is attached to the learner, who is actually an actor. Subjects are assigned to the role of teacher. The teacher is to ask the learner a series of questions, and deliver a shock every time the learner gets the answer wrong. Each shock must be of greater intensity than the last. In this experiment, the learner continually provides incorrect answers, and asks to be let out. However, the experimenter directs the subject to continue giving higher shocks. The learner eventually refuses to continue, appears to have a heart attack and stops responding altogether. Despite the obvious wrongness of electrocuting someone who has had a heart attack, Milgram found that 65% of subjects continued administering shocks at the request of the experimenter.

The results of these studies are very surprising to most people for a number of reasons. First, it seems very obvious what the right thing to do is. The subjects clearly ought to disobey the experimenter. Second, nothing appears to be impeding the subjects from doing the right thing. They are not coerced or compelled, and there seems to be very little cost to doing the right thing. Third, it seems that if asked ahead of time how they react, most subjects would report they would disobey, and would likely believe this too. As a result, the subjects' behaviour conflicts with our folk predictions of how most people would act in these settings. Milgram famously surveyed psychologists prior to conducting these studies, and found they predicted only 1–2% of people would continue obeying until the end (Milgram 1974).

While much philosophical analysis of the situationist experiments has focussed on why certain factors can have such substantial effects on our moral behaviour, my area of interest is the *prevalence* of this immoral behaviour. Specifically, I believe that the fact that such high proportions of people act immorally in some settings at all is philosophically important for thinking about whether we it is appropriate for us to hold these people responsible.

3. Obedience to authority and moral responsibility

Many people are somewhat reluctant to blame the obedient subjects in Milgram's experiment upon hearing about this experiment for the first time. Given the obedient subjects' behaviour conflicts with what we think is the obvious course of action, is influenced by morally irrelevant features, and is seen in such a high proportion of subjects, this behaviour appears odd to say the least. This can lead people to think subjects' rational capacities are impeded in some way. As Nelkin (2005) puts it, 'the subjects seem to be acting for bad reasons, or at least not acting for good reasons, and they seem stuck doing so' (p. 199).

I believe part of what explains this intuition is the realisation that we, and many people who we think are generally good-hearted, would have obeyed too. The realisation that most of us can be induced to behave immorally in these settings sits uncomfortably with our beliefs that obeying is clearly morally wrong, and that we are generally good people who wouldn't harm strangers in settings where avoiding wrongdoing is relatively easy. It is therefore tempting to think that if we in fact would have acted so impermissibly, it must be because something compromised our agency or deliberative capacities. This prompts a search for evidence that the obedient subjects' capacities were impeded in some way.

Fortunately, a number of philosophers have already examined whether subjects who act immorally in the situationist experiments have reduced culpability due to their capacities being impeded in some way. Though the obedient subjects' behaviour is odd, the consensus amongst philosophers is that subjects are not excused for their behaviour. Brink (2013) argues that since it is likely the subjects in Milgram's experiment would have disobeyed given more time to deliberate or greater incentives, they had the capacity to disobey but simply failed to exercise this capacity. McKenna and Warmke (2017) argue that though the combined situationist literature suggests we have less capacity to avoid wrongdoing than we like to think, subjects in these settings are still reasons-responsive to the degree required to be considered morally responsible agents. Levy (2014, 2015) argues that even if situationist factors can change which action an agent performs, they can't make subjects act in such a way that conflicts with their values entirely. Because subjects are aware of what they are doing, and have the capacity to do otherwise, they are responsible for their actions. Though unusual and distressing, most philosophers think that nothing relevant to our culpability assessments distinguishes the choices these subjects make from decisions they make

in other settings. The stress of the situation may mean subjects are less blameworthy than they would have been had they gleefully administered the shocks, but it does not exculpate them. This conclusion seems to entail that our initial reluctance to blame is therefore unwarranted. If nothing affected subjects' rational capacities, they are morally responsible and thus blameworthy for their actions.

Though I agree with these philosophers that the obedient subjects are blameworthy for their actions, I believe that many people's hesitation in blaming the obedient subjects is nevertheless tracking something important. This is because the fact that an agent is blameworthy does not always mean the agent is an appropriate target of our blame. In addition to wrongdoers needing to first meet certain conditions for blame to be appropriate, blam*ers* must also meet certain conditions for their blame to be appropriate. Philosophers describe agents who don't meet these conditions as lacking the *standing* to blame.[1]

The most common condition thought to undermine standing to blame is hypocrisy. Or, rather, agents who culpably commit a certain wrong are thought to lack the standing to blame others who commit similar wrongs, and become hypocrites if they in fact blame others. Many people object to hypocritical blame, and multiple philosophers have recently argued that hypocrites lack the moral standing to blame wrongdoers for actions the hypocrite is also guilty of performing (Coates and Tognazzini 2013; Friedman 2013; Roadevin 2018; Todd 2019; Wallace 2010). Hypocritical blaming is often thought to evince some kind of moral fault, even if the blamee is indeed culpable. If Sean and Alison both cheat on their spouses, Alison is entitled to judge that Sean has done something wrong. Alison is entitled to think Sean has a callous character, and inform Sean that there are strong moral reasons to feel some remorse. Blame, however, is off the table. If Alison were to openly rebuke Sean, Sean – as morally flawed as he is – is within his rights to say 'I know what I did was wrong, but who are *you* to blame *me*?' This response acknowledges that the speaker has committed a wrong and

[1]I follow Fritz and Miller (2018) in thinking of standing to blame as a kind of defeasible entitlement or right, which plays an important role in the ethics of blame. Philosophers writing on moral responsibility recognise that the question of whether blame is appropriate is different to the question of whether blame is justified, all-things-considered. If you possess standing to blame a culpable wrongdoer, but blaming them would cause them to go into a deep depression, you ought not to blame them. Blame would be appropriate, but not justified. If you lack the standing to blame someone (either due to facts about yourself or because they are excused) but blaming them would avoid some terrible outcome, then you ought to blame them. Blame would be inappropriate, but justified. Questions of standing regard the appropriateness of blaming.

that certain responses from the moral community are appropriate, but denies that the hypocrite is in any position to blame.

If hypocrites lack the standing to blame in virtue of something about their prior wrongdoing, I believe we have a ready explanation for why many people are reluctant to blame the obedient subjects in Milgram's experiment. If we can lack the standing to blame people for things we are equally blameworthy of, it is only a small step from there to think we can also lack the standing to blame people for wrongs we *would* have committed if we had been placed in that situation. This would entail that most of us lack the standing to blame subjects who obey in Milgram's experiment, because we likely would have obeyed too.

This argument has a number of steps which need more unpacking and defence however, so let me sketch out the remainder of this paper's argument. First, I present Todd's (2019) account of standing to blame, and argue it entails that many subjunctive wrongdoers also lack the standing to blame. Combining this with findings from Milgram's experiments and its many replications leads to the conclusion that most of us lack the standing to blame subjects who act wrongly in these settings.

4. Blame, standing and hypocrisy

I believe the strongest account of standing to blame comes from Todd (2019).[2] Through careful consideration of a number of thought

[2]Let me mention some objections to other accounts of hypocrisy, for readers familiar with the literature. Fritz and Miller (2018) argue that hypocrites have a 'differential blaming disposition' which implicitly rejects the impartiality of morality and thereby denies the equality of persons. Todd (2019) shows their account implausibly entails that agents with dispositions to inconsistently blame wrongdoers for particular wrongs can lose their standing to blame anyone for similar wrongs altogether. For example, it would be inappropriate for citizens of Western nations to blame terrorists for an attack in France if those citizens failed to blame terrorists for an attack in Turkey some months earlier.

Wallace (2010) argues that when we blame, we undertake a commitment to a principle of equality. Hypocrites elevate their interests above the wrongdoer's, denying the equal standing of persons that is 'practically constitutive of morality' (p. 330) in the first place. I reject Wallace's account because it entails that anyone who blames without self-scrutinising is open to the same objection we make against hypocrites, even if the blamer is a moral saint.

Roadevin (2018) argues that hypocrites treat the blamed unfairly, by demanding an apology without providing one themselves. This account can't accommodate our intuitions regarding *Diet* and *Mistaken Blame* (described below), and the claim that expressed blame seeks an apology is dubious as we can often appropriately blame people after they've apologized.

Isserow and Klein (2017) argue that hypocrites undermine their claim to moral authority in virtue of their actions. Because this is an account of hypocritical moral *advice*, rather than hypocritical blame, I don't have any strong objections to the overall account. But regarding blame, the authors don't identify *why* hypocrites lack an entitlement to claim moral authority, only briefly alluding to Cohen's (2006) suggestion that some a 'disabling fact' means hypocrites literally cannot perform the speech act of condemning. In any case, Isserow and Klein's account seems compatible with my conclusion: if in blaming the obedient subjects, we are making a claim to practical moral authority, the fact that we would not guide our own actions by this advice undermines our claim.

experiments related to hypocrisy, Todd convincingly argues that agents lack the standing to blame a wrongdoer when they are not sufficiently committed to the kinds of values that would condemn the wrongdoer's actions.[3] Todd starts his argument by first noting that we often think that agents lack standing to blame even if they haven't performed the same wrong as the person they are blaming. Consider this case, originally from King (2015):

> *Diet*: Suppose Charlie knows that Linus, who has a weakness for sweets, is trying to lose weight. Nevertheless, he takes Linus to a place for dinner that he knows is located next to an incredible ice cream shop. Quite predictably, after dinner Linus visits the shop next door and has some ice cream. King (2015, 5–6).

Linus is to blame for violating his commitment to his diet, in that it would be appropriate for his family and friends to blame and hold him responsible for this. But even though Linus is fully responsible for his wrongdoing, Charlie is also partly responsible, and not in a position to appropriately blame Linus.

Second, agents who do commit the same action as the blamed can escape the charge of hypocrisy and retain standing to blame if they perform that action for justifiable reasons:

> *Nazis*: Jonas is a Nazi commander working in a death camp [but] is secretly opposed to the Nazi regime. He thus does everything within his power to save the lives of as many prisoners as possible, consistent, of course, with maintaining his position as a committed Nazi; Jonas (correctly) reasons that he can do much more good secretly sabotaging the Nazi efforts as a trusted commander than he could by open defiance. Jonas hears rumors of an escape. In order to keep appearances, he must order someone to investigate the fence. Jonas thus orders Thomas to investigate the fence and sound the alarm should he see anyone attempting escape. Jonas chose Thomas for this task because he (blamelessly, though incorrectly) thought that, of all the people he might choose, Thomas would be the most likely to have mercy and not sound the alarm should he actually find prisoners escaping, and instead report back that there was nothing to the rumors. Instead, however, Thomas discovers the escaping prisoners, sounds the alarm, and the prisoners are caught and executed. Todd (2019, 354).

Were Jonas to blame Thomas, this blame would be appropriate, and Thomas has no grounds on which to object even though Jonas ordered

[3]Rossi (2018) also argues for a similar claim, taking agents to be hypocritical when they are lack sufficient committed to a particular norm, good or ideal, but are disposed to communicate that they are committed to it. His account would be compatible with my argument in this paper. For simplicity, I focus on Todd's account. See Fritz and Miller (2019) for some objections to Rossi's account.

Thomas to act as he did. Jonas only gave the orders to avoid detection, and his motives make him an inappropriate target of our ordinary pejorative charge of hypocrisy.

Third, agents who have lost their standing to blame can regain this standing by now being committed to the relevant values that condemn their prior wrongdoing:

> *Abandoned*: Paul, when he is 25, abandons his wife and newborn infant to go live an unencumbered life touring around the world and "find himself", and that his wife and child suffer terribly as a result. Suppose further that later in life, at the age of 45, he comes to realize the virtues and obligations of fidelity and becomes a committed husband and father to his wife and child, with whom he is somehow reconciled. He now holds values, very deeply, which condemn his past actions. Now suppose that finally hearing of another 25-year-old who has abandoned his wife and child to explore the world just as he once did he becomes indignant with this other person. Todd (2019, 357).

Even though Paul's past actions are just as blameworthy the actions of the 25-year-old he is now blaming, and even though Paul actually lost his standing (unlike Jonas), our judgments regarding standing are sensitive to how agents are presently constituted.

Fourth, people can lack the standing to blame agents they mistakenly believe exist:

> *Mistaken Blame*: Samuel is a shop owner who has himself never tried to shoplift – but not for want of wishing to do so. Samuel simply lacks the means to try – a fact he regrets. He simply has no vehicle to travel to other stores, and there are no suitable targets within walking distance. (If there were, Samuel explains, he would try to shoplift from those stores.) One day, Samuel gets a call: someone has been shoplifting from his own store. On hearing this, Samuel burns with rage and incredulity and a desire to confront the criminal. Todd (2019, 362).

What is important to note is that Samuel's hypocritical blame doesn't in fact wrong anyone or treat anyone unfairly. Nevertheless, he isn't entitled to express blame in virtue of his prior wrongdoing, and third parties would point this out to him. It would be appropriate to respond with 'Who are *you* to blame them?' or 'Look who's talking' to him blaming.

Todd's conclusion is that agents lack the standing to blame a wrongdoer when they are not sufficiently committed to the values that would condemn the wrongdoer's actions. This account of standing to blame is an improvement on other accounts in that it is simple, able to explain many of our intuitions across a range of cases, and doesn't seem to face any obvious counterexamples. Note it has the interesting consequence that prior wrongdoing doesn't directly undermine standing to blame.

Rather, a hypocrite's prior wrongdoing is strong evidence that they are not sufficiently committed to the values which would condemn the other person's wrongdoing, and this lack of commitment undermines their standing to blame.

Todd's account has three weaknesses. The first is that Todd doesn't yet have an answer to why commitment to the relevant values is necessary for standing to blame, something he acknowledges (p. 371). Though providing a full account would be outside the scope of this paper, and nothing in this paper turns on having a specific explanation for why commitment to the relevant values is necessary for standing to blame, let me gesture towards a tentative answer. I believe that part of the answer relies on the common thought that blame expresses demands (Darwall 2006; Watson 1996). Since demands are things which agents must have authority to make, it could be that sufficient commitment to relevant values is necessary to have the moral authority to make certain kinds of demands on others.[4] Philosophers who reject the claim that blame expresses demands (e.g. Macnamara 2013) can still accept the remainder of my argument, if they agree with me that Todd has correctly identified the relevant instances in which agents lack standing to blame. For the remainder of this paper, I will assume that Todd's account is correct.

The remaining weaknesses of Todd's account are that he never explicitly investigates how widely or narrowly we should think of these values, what counts as 'sufficient commitment', nor how to decide whether an agent is in fact sufficiently committed to a moral value. To improve his argument and overcome this gap in the literature, I will take up these questions in §6. For now, I follow Todd in noting that commitment to a value consists in some endorsement of that value, and some degree of motivation to comply with that value.

To avoid unwieldy sentences or possible ambiguity, let me stipulate that for the remainder of this paper that when I say an agent 'lacks the standing to blame' I mean they lack the standing to blame people for wrongs which would be condemned by values that agent is not sufficiently committed to. It is not the case that hypocrites lack the standing to blame any wrongdoer whatsoever. Let me also stipulate that the moral value which would condemn the wrong action or omission under consideration is what I will call 'V'. Agents who have some minor endorsement of the relevant moral value, but not enough to gain standing to

[4]I argue for such an account in Piovarchy (2020).

blame, will lack sufficient commitment to V. Since I accept Todd's account, to assess whether an agent has the standing to blame, we need to assess whether they are sufficiently committed to V.

5. Subjunctive hypocrisy and moral luck

If one believes that hypocrites often lack the standing to blame in virtue of their prior wrongdoing, it is then only a small step to think agents who *would* count as hypocrites, but for chance, will also lack the standing to blame. Suppose we have a wrongdoer who is culpable for ϕ-ing. Call agents who would have culpably ϕ-ed in similar circumstances subjunctive wrongdoers. If our subjunctive wrongdoer were to blame an agent who in fact ϕ-ed, we might be tempted to label the subjunctive wrongdoer a subjunctive hypocrite. Little has been written on the relationship between standing to blame and what I am calling subjunctive hypocrisy, but Todd's account can easily identify which kinds of subjunctive hypocrites lack standing to blame.[5]

Agents who ϕ but blame others for ϕ-ing only lack the standing to blame when that agent lacks a sufficient commitment to V. This entails that agents who *would* have ϕ-ed due to a lack of sufficient commitment V, but who haven't ϕ-ed simply due to not being in the relevant setting, will also lack the standing to blame. We don't necessarily need to investigate whether an agent would have ϕ-ed to determine whether they lack standing to blame, as we can just investigate whether they possess sufficient commitment to V. But the fact an agent would ϕ can be great evidence they lack sufficient commitment to V.

At first glance, thinking about agents who would have done the wrong thing but haven't simply due to chance seems to introduce questions of moral luck. Philosophers who claim that moral luck exists believe that an agent's level of responsibility and blameworthiness can be can be affected by factors beyond her control (Nelkin 2014). But Todd's account actually doesn't depend on whether moral luck of any kind affects responsibility or blameworthiness. Agents lack standing to blame when they lack sufficient commitment to V, and this is not sensitive to questions of whether factors beyond one's control can affect one's level of culpability.[6]

[5]Rivera López (2017) also concludes that circumstantial moral luck doesn't affect standing to blame, but he doesn't have any unified explanation for what undermines standing to blame, or what makes one wrong 'relevantly similar' to another. By identifying an agent's commitment to values as the key factor, Todd's account seems to unite a number of factors that Rivera-López implicitly took to be important.

[6]In some cases, luck in the way things turn out might cause an agent to become committed to the relevant values and thereby cause them to regain standing. But this is distinct from moral luck, the thesis that luck can affect one's level of responsibility or blameworthiness.

For example, if two assassins try to murder the same person simultaneously, but a chance gust of wind causes one assassin's bullet to miss, luck has affected the outcomes of their actions. This luck hasn't had any effect on either assassin's commitment to the relevant values, however. Similarly, if two agents would act wrongly in a certain situation due to both lacking sufficient commitment to V, but only one agent gets placed in that situation and in fact acts wrongly, luck has affected whether each agent committed any wrongdoing. It is an open question as to whether luck has thereby affected whether these agents should be thought of as equally blameworthy. However, luck hasn't affected whether each agent presently is sufficiently committed to V, and so luck hasn't affected either agent's standing to blame.

Luck in how agents are constituted can change one's standing to blame. The environment that an agent is raised in, their genetics and thus their constitution later in life are all significantly affected by chance. But there is no reason to think that this unfairly affects standing to blame, and so I am happy to accept this consequence of Todd's account. The fact that you would have done the same thing, and thus would have lacked standing if you had another person's genetics and upbringing, says nothing about how you are *actually* constituted. Additionally, this line of thinking is subject to a well-known objection. The thought that you would have done otherwise if you had a different history and constitution simply amounts to the thought that 'I would have done the same thing if I was a different person', and it is hard to make sense of how exactly 'I' could be a different person (Nagel 1979).

Instead, the thought that we would have done the same thing if we had been a different person seems more relevant to the question of whether the other wrongdoer had a fair opportunity to avoid wrongdoing, or whether it was reasonable to expect them to do otherwise. This is a question about whether the other person is an appropriate target of blame at all, not whether it is appropriate for a blamer to blame given certain facts about the blamer. And this is a distinct question for philosophers working on problems of constitutive moral luck or free will and determinism, which we can set aside.

6. Scope, evidence, and commitments

If insufficient commitment to V undermines standing to blame, and if subjunctive wrongdoing can be strong evidence of insufficient commitment to V, the situationist experiments have important implications for our

blaming practices and the moral community. To the extent that Darley and Latane's (1968) study is replicable in today's population, we should infer that 69% of us would fall prey to the bystander effect and fail to give aid in settings similar to those in the experiment. We should also conclude from Darley and Batson (1973) that only 10% of us would give aid to an unconscious person if we found them while late to deliver a speech.

The most distressing result is Milgram's obedience to authority experiments, and it is these experiments for which our preceding account of standing to blame has the clearest philosophical implications. Given this experiment has been well replicated in a number of cultures across a number of decades with very similar results (Blass 1999a; Burger 2009; Dolinski et al. 2017; Edwards et al. 1969; Shanab and Yahya 1978), we should believe that roughly 65% of us would have continued administering shocks all the way to the end of the scale had we been subjects. We would have continued administering shocks even after the learner asked to be let out, was audibly in pain, stopped responding, and appeared to have had a heart attack.

It is important to note that although it is commonly reported that 35% of subjects 'disobeyed', most of these subjects still went far beyond what is morally acceptable. Over 80% of subjects continued administering shocks after the learner demanded to be let out and complained about his heart. Had the experiment been real, 80% of subjects would have given the learner a heart attack and still not left the room to check on him.

However, the fact that most of us would obey in this setting doesn't yet show that we lack the standing to blame subjects who in fact obeyed in this setting. To show that 80% of us lack the standing to blame the obedient subjects in Milgram's experiments, I need to show that those of us who would obey lack sufficient commitment to V. As noted earlier, there are three components of Todd's account that we need to provide answers for: how widely or narrowly we should think of V, what counts as sufficient commitment to V, and how to determine whether a particular agent is sufficiently committed to V.

How narrow we should understand V is an important question. If we construe it too narrowly, an assassin who strangles her victims will have the standing to blame a different assassin who shoots his victims, since condemning strangling is different to condemning shooting. On the other hand, if we construe it too widely, adulterers will lack the standing to blame assassins since it could be argued both are not sufficiently committed to doing-the-right-thing.

Worries about scope do not apply to the case at hand. This is because the subjunctive wrongdoing under consideration is the *exact same* kind of wrongdoing as that committed by the obedient subjects. Whichever values condemn the wrongdoing of the obedient actions today will be the same values which would condemn the wrongdoing of most blamers were they to end up in a Milgram-style experiment. There may be some exceptions in which agents obey for completely different reasons, in such a way that shows they are sufficiently committed to V. But such cases would be rare, and it seems hard for them to arise without it being the case that the subjunctive wrongdoer would be blameworthy in a different way. Suppose, for instance, that one subject would obey only because she had a deep phobia of scientists in white coats. This person's obedience might come about in such a way that renders their wrongdoing different to the wrongdoing of subjects without this phobia. But in such a case it seems this person had a reduced capacity to disobey, and this could affect our degree or manner of blame towards them.

Since these studies are well-replicated in a variety of populations, and those who obey do so in very similar ways, most subjunctive wrongdoers would be blameworthy in the same way as the obedient subjects. They would thus be condemned by the same kinds of values those wrongdoers would invoke when condemning the obedient subjects. This means that if it can be shown that the obedient subjects lacked sufficient commitment to V, it will follow that the majority of us who would obey in the experiment will also lack sufficient commitment to V and thus lack the standing to blame the obedient subjects.

Our questions now are what counts as sufficient commitment to V, and how to determine when agents possess sufficient commitment to V. Even if we're uncertain on the answer to the former question, it could be argued that there are two intuitively good sources of evidence that the obedient subjects did possess sufficient commitment to V, because these are the kinds of evidence we generally use when deciding if agents value something. We generally expect agents who value something to act in ways consistent with valuing that thing across a range of domains. Someone who values friendship will generally wish their friends happy birthday and visit them when they are sick. We also expect agents who value something to experience certain affective responses indicative of a commitment to that value, such as caring that the value is adhered to and being distressed when it is not. If someone values their friendship with a person, they will care about how that person is doing, and experience appropriate affective responses in response to the other person's well-being.

The obedient subjects seem to meet both these criteria. They presumably had strong track records of not harming people. They would have previously had many opportunities to harm people, but successfully refrained from doing so. Additionally, most subjects were distressed during the experiment and attempted to help the learner.[7] They clearly had a strong desire to not harm the learner, and would probably report having a strong commitment to not harming strangers.

I argue that this evidence is not enough to establish the obedient subjects possessed a sufficient commitment to V. Appealing to the agents' distress as evidence they possessed sufficient commitment to V relies on a common but mistaken standard by which we evaluate people's degree of commitment. Suppose Sally volunteers at the library every week instead of using her spare time to watch Netflix, and seems to experience the relevant appropriate responses to her patrons or hearing that the library has received extra funding. We would usually conclude Sally is committed to, and values, volunteering at the library. But this method is problematic because it only assesses how committed an agent is relative to their other preferences. To show what I mean, imagine we now find out Sally would stop helping at the library permanently if she was offered $10 to quit, or if her boss asked her to come one minute earlier from now on, or if a new show was added to Netflix. We would be inclined to think that she was in fact never very committed to helping at the library after all. Helping at the library had simply been what she preferred to do out of a narrow set of options.

Because we are assessing the obedient subjects' commitment to *moral* values, we can't ascribe sufficient commitment to those values simply because the subjects had a desire to act in a way which was consistent with those values, when acting immorally. The subjects' commitment to those values might still be very weak on some absolute or objective standard. Their commitment to those values certainly seems very weak relative to the level of commitment they ought to have had.

One might think that the subjects' histories are evidence that they were committed to V on some absolute measure of commitment. But it is important to note that while subjects' histories are good evidence that they don't intrinsically value harming people, we are blaming the subjects for not refusing to administer shocks when directed to by an authority figure. The fact that the obedient subjects had a strong history of not

[7]With some exceptions: Paul Batta seemed to display 'total indifference' towards the learner, and 'a submissive and courteous' manner to the experimenter (Milgram 1974, 46).

harming people doesn't show they were strongly committed the values which would condemn their wrong actions, as they simply hadn't been in situations where there was strong social pressure to obey someone in a lab coat.[8] If the learner had in fact been harmed, they would appropriately reproach the subject for not being more motivated to act in accordance with their better judgment, and for not having the fortitude to stand up to the experimenter. It is very possible for someone to feel distressed about acting wrongly, while still choosing to act wrongly due to insufficient commitment to the kinds of values which would condemn that wrong action.

7. Diagnostic acts and capacities

Whether the obedient subjects possessed sufficient commitment to V remains an open question. We thus need some means of evaluating when an agent has sufficient commitment to V. This is difficult, because our assessments of an agent's commitments are often very context-sensitive. The strongest evidence of whether an agent is sufficiently committed to V isn't just that we observe a lot of behaviours consistent with V. It's that they regularly display V-consistent behaviour in a range of V-eliciting circumstances. Likewise, a failure to display these behaviours in these circumstances is our strongest evidence the agent lacks sufficient commitment to V. This is evidence we currently don't have, as the obedient subjects haven't been placed in a range of relevantly similar settings. But I don't think we need to start speculating about how the obedient subjects would act in such settings.

This is because for many values, a single instance of severe and culpable wrongdoing can be *diagnostic* that an agent is not sufficient commitment to that value. While someone who eats one unhealthy meal can still be considered sufficiently committed to their diet, someone who murders once cannot (at that time) be considered sufficiently committed to valuing innocent human life.[9] A single act of wrongful intentional

[8]This is similar to Doris's (2002) argument that in order to be warranted in ascribing global character traits to an agent, we need to observe trait-consistent behavior in a range of trait-eliciting circumstances consistently more often than we would expect by chance. My point is that the obedient subjects simply haven't been in a range of V-eliciting circumstances.

[9]Badhwar (2009) puts the point nicely when assessing whether the obedient subjects possess certain virtues: 'It is true that the failure to act virtuously in one situation does not, typically, reveal a lack of the relevant virtues, whether global or domain-specific. But all situations are not equal. Suppose a man murders his wife simply to inherit her property and marry his wealthy mistress. Must we wait for him to murder this new wife in order to inherit her property and marry his new mistress before we can justifiably conclude that his respect for human life is not very deep? Evidently not; the situation

murder is not merely in tension with valuing human life, it is entirely incompatible with having this value, even if you were distressed while carrying it out and felt regret afterwards. Recall that had the experiment been real, the obedient subjects would have been guilty of manslaughter, and 80% of subjects would have been guilty of shocking someone until he had a heart attack. They didn't call the ambulance, and they didn't leave the room to check on the learner.

Though the subjects didn't show any malice in their actions, they nevertheless committed severe wrongdoing. In cases of severe and avoidable wrongdoing, we are entitled to a default assumption that the wrongdoer was not sufficiently committed to the moral value which would condemn that wrong. We may later gain further evidence that there were extenuating circumstances, or that they have since atoned and were now sufficiently committed. But these cases are the exception, not the rule. Arguing that the we possess standing to blame requires an explanation of how subjects could have carried out their actions in such a way that is compatible with possessing sufficient commitment to V.

To be sure, the subjects were in an unusual setting, and we might think that it was psychologically difficult in some sense to disobey. But when philosophers and lawyers discuss the difficulty of avoiding wrongdoing as a mitigating factor, they have in mind difficulties which directly impede the exercise of an agents' capacities. They have in mind cases where agents are either coerced into wrongdoing, or where avoid wrongdoing carries with it significant cost (Brink and Nelkin 2013). The difficulty of acting against a strong desire usually isn't taken to be mitigating, unless it can be shown that this desire in some way reduced the agent's capacity to avoid wrongdoing, as in cases of compulsion or addiction. These kinds of mitigating factors were not present for the subjects in Milgram's experiments. It is also worth emphasising just how low of a bar subjects had to clear to qualify as having 'disobeyed'. Subjects didn't have to outright refuse or challenge the experimenter. They merely had to voice some kind of concern or question four times in a row. At this point, the experiment was terminated *by the experimenters*.[10] As noted in §4, numerous

is "diagnostic" (to adapt Doris's term for my own purposes) of a lack of respect for human life. The situation in the Milgram experiment is similarly diagnostic. The experiments offer a clear choice between good and evil: both common sense and the direct evidence of their senses tell the subjects that the shocks are painful as well as dangerous, and basic moral decency tells them that it is wrong to harm an innocent person against his will for a trivial purpose. Moreover, the choice facing them is an easy one, since disobedience poses no threat to their lives or happiness' (p. 263).

[10]There appears to be some confusion regarding the accuracy of this point. Milgram originally reported that the experimenter used a 'firm but not impolite' (1963, p. 374) voice and that the experiment was ended if subjects questioned the experimenter four times. Perry's *Behind The Shock Machine*

philosophers have convincingly argued subjects had the capacity and opportunity to disobey.[11] The subjects simply didn't exercise this capacity. They were aware their actions were wrong, but they didn't follow their better judgment due to a fear of embarrassment or desire to comply with perceived social norms (Sabini and Silver 2005).

Given this severe wrongdoing and the low cost of disobeying, the fact that subjects had not harmed people previously and felt distressed during the experiment is not enough to ascribe to them a strong commitment to the kinds of values which would condemn their wrongdoing. I thus believe that the obedient subjects were not sufficiently committed to the values which would condemn their wrongdoing prior to and during the experiment.[12] If they were sufficiently committed, they would have been sufficiently motivated to disobey.

Since the obedient subjects in these experiments are ordinary agents just like us, and since there are many replications showing that many different kinds of people obey the experimenter in this setting, this means that approximately 80% of us lack the standing to blame the obedient subjects. Of course, it is difficult to identify whether any particular one of us falls into this group, and whether being well-informed about situationist experiments might make us more likely to disobey. But all else being equal, it is inappropriate for approximately 80% of us to blame the obedient subjects for their wrongdoing.

8. Objections

One possible objection to the claim that standing to blame requires sufficient commitment to V is that it we don't often make any 'Who are

(2013) reveals that the 'experimenter', an actor, initially followed this script. But as the trials went on he strayed from the script, requiring more and more refusals to qualify as disobedient (p. 115–119). Perry notes he was particularly aggressive with female subjects, arguing this qualified as coercion. I agree this would give us cause to worry about the validity of the experiment, were it not for the numerous replications in which this coercion is absent. These replications were far more rigorous in terminating the experiment whenever subjects 'expressed either verbally or nonverbally a reluctance to continue the study' four times (Blass 1999b; Burger 2009, 7; Doliński et al. 2017; Miller 1986).

[11] Real self and attributionist accounts of moral responsibility (e.g. Talbert 2009) argue agents are blameworthy because their actions are attributable to them, and so may not grant that the subjects had the capacity to do otherwise. I don't think this is a problem, because locating responsibility in how the agent is constituted seems rather close to locating responsibility in the way the agents' values and motivational set are arranged. If the agent possessed sufficient commitment to V but still committed wrongdoing, then since sufficient commitment to V entails acceptance of a value and motivation to comply with that value, attributionists may be inclined to say that their wrongdoing is going to be explained by factors showing that the wrong wasn't attributable to the agent. This would make the obedient subjects excused, thus rejecting our starting premise that these subjects are culpable.

[12] They could plausibly gain sufficient commitment to V after the experiment, however.

you to blame *me*?' response to people who are subjunctive wrongdoers. Getting exercised about subjunctive wrongdoers who blame us may happen sometimes, but it doesn't seem like a regular feature of our moral practices, and therefore my account seems rather revisionary. While I agree with this observation, my account is not revisionary as there is a good explanation for why we don't commonly have these responses. It's not that my account conflicts with our moral practices, rather, it's that we are rarely in settings where we can recognise that someone lacks the standing to blame due to a lack of sufficient commitment to V. This is due to the simple fact that we are epistemically limited in a number of ways, and our primary source of evidence for people's commitments is their behaviour.

We cannot see the future behaviour of others or ourselves, whereas we can have directly observed people committing wrongs in the past. We thus have much more evidence and greater ability to identify actual wrongdoers who lack the standing to blame, even though there exist more subjunctive wrongdoers who also lack standing to blame. It is also very difficult to assess whether we would violate a moral principle in certain circumstances, because asking and answering that question can change how we would behave in those circumstances. Reflecting on your own dispositions to wrongdoing can motivate you to avoid acting in that way, thereby changing your disposition and thus whether you are a subjunctive wrongdoer. Likewise, presently judging someone to be a subjunctive wrongdoer will often motivate them to avoid acting in the way you predict they would act. They will be eager to prove you wrong, and want to report that they would have done the right thing all along. Because of these epistemic limitations, it is very difficult to make assessments regarding who possesses sufficient commitment to V and adjust our blaming practices accordingly, even if it is true that subjunctive wrongdoers who lack sufficient commitment to V lack the standing to blame.

Trying to overcome these epistemic difficulties would come at a considerable cost. Because it is difficult to assess whether an agent is a subjunctive wrongdoer who lacks sufficient commitment to V, reaching any shared conclusion on which members of the moral community have standing to blame would often take a lot of time. Additionally, making this a regular feature of our practices could give culpable wrongdoers a strong incentive to begin questioning the standing of blamers in an attempt to avoid being held responsible. As a result, our need to avoid perverse incentives, our epistemic limitations, and our desire to make sure wrongdoers are held to account leads to a reasonable default assumption

that members of the moral community have standing to blame. These factors have meant that philosophers haven't had much motivation to analyse the relationship between subjunctive wrongdoing and blame, even though the claim that subjunctive wrongdoers who lack sufficient commitment to V lack the standing to blame is consistent with our moral practices.

Another objection is that if subjunctive wrongdoers who lack sufficient commitment to V lack standing to blame, this leads to a number of undesirable consequences. We are forced to condone the actions of wrongdoers, or at least aren't entitled to complain. This account removes incentives against wrongdoing, and could make us unable to engage in moral address in settings where this would be effective. Additionally, this conclusion supports a world where wrongdoers, when being blamed, have license to start bringing up faults that the blamer has committed, instead of focussing on their own faults and making amends.

I don't think any of these consequences follow from my position. First, nothing I have said prohibits other reactions to wrongdoing, or other forms of engagement with wrongdoers. One is still free to try to engage in moral education, to convince the wrongdoer that they've done something wrong and need to make reparations, or to talk about one's own experience acting wrongly and why one regrets it.

It is also worth emphasising that agents can regain standing to blame by becoming the type of person who is now sufficiently committed to the relevant values. I agree that anyone who focuses on others' faults instead of their own when being blamed displays a serious moral fault. I do not intend to promote such a mentality. Instead, my aim is to encourage would-be blamers to take a step back and check that their blame stems from a genuine commitment to moral values, rather than being a knee-jerk reaction stemming from problematic motivations such as moral grandstanding (Tosi and Warmke 2016).

Recognising our own subjunctive shortcomings also has benefits. When we are in a position to blame others, it is easy to think that recognising an act is wrong should be enough to motivate people to avoid committing it. But when we realise that we would have done the same, we can appreciate that wrongness alone is not sufficient to motivate people to act well and so other solutions are needed. This makes us much more likely to begin thinking of more effective ways to prevent those settings from developing altogether. We are able to use our understanding of what would cause us to become wrongdoers to design systems with checks and balances, and look more broadly for preventative measures. Instead

of just telling people in positions of power (e.g. experimenters) to be ethical, we can design environments in ways that help foster moral behaviour (Russell 2009).

9. Some notes on generalisability

A relevant question here concerns whether we lack the standing to blame subjects who act wrongly in other situationist experiments. Since there are hundreds of situationist experiments, and each identifies a different factor which can increase our chances of wrongdoing, this might make it seem like we lack the standing to blame far more wrongdoers than is ordinarily assumed.

I believe that most of the situationist experiments don't directly show we lack the standing to blame subjects who act wrongly in these settings. This is because the wrongs under consideration in many experiments are very minor, such as failing to help pick up some papers (Isen and Levin 1972) or lying about one's score on a test (Gino, Norton, and Ariely 2010). While subjects' behaviour falls below the standard we think it ought to have been, it is not grossly lower. These single instances of behaviour which are inconsistent with a moral value are not enough to demonstrate those agents are not sufficiently committed to that value. Subjects (and we) may in fact lack the relevant moral values which would condemn these wrongs, but these situationist experiments are not proof of this.

It is harder to assess the commitments of subjects who acted wrongly in other situationist experiments featuring strong social pressures. These experiments show subjects failing to help an unconscious person (Darley and Batson 1973) or failing to call for help after hearing someone have a seizure (Latané and Darley 1970). Though these actions are less wrong than those of Milgram's obedient subjects, they still seem like significant wrongs, and the kind of things for which strong blame is usually appropriate. I am inclined to say that we do not have enough evidence to settle the matter. We can only say that if we would fail to help in this settings, and if that failure to help is attributable to insufficient commitment to V, then we would lack the standing to blame those subjects who in fact did not help in these settings.

My indeterminate assessment of subject behaviour in these experiments may seem unsatisfying, but our ability to assess peoples' degree of commitment to moral values is inherently limited by the kinds of evidence we have available. Assessing people's degree of commitment to

certain values is difficult because most people simply don't find themselves in the kinds of settings which act as the clearest evidence. Even when people are in those settings and we can observe how they in fact act, we don't know how those people would act in relevant counterfactual settings. And trying to assess how people would act in these counterfactual situations via introspection is likely to be unreliable or self-serving. One of the key lessons from the situationist experiments is that we are generally very bad at predicting how we will act in certain situations, and at identifying what sorts of factors influence our behaviour.

One final point on the generalisability of these studies is worth noting. The situationist experiments are suggestive of another way in which fewer of us lack the standing to blame than is commonly assumed. This comes from the literature on the consistency of people's behaviour across normatively similar but otherwise distinct situations. For example, while we usually think that honest people won't cheat on exams, steal candy bars or lie to their friends, Hartshorne, May, and Maller (1929) notably found that any observed instance of honesty in one setting is a very poor predictor of these other kinds of honesty in other settings ($r = .23$). Mischel and Peake (1982) and Chaplin and Goldberg (1984) also found very low correlations when assessing people's behaviour across a range of situations. Though they weren't specifically interested in the kinds of behaviour which enable us to discern an agent's commitment to certain values, this is nonetheless worrying because we expect agents who are committed to certain values to consistently act in accordance with that value across a range of relevant circumstances.

This lack of consistency matters, because agents who successfully avoid severe wrongdoing in one setting, and to whom we are inclined to attribute sufficient commitment to V, may commit a similar and severe wrongdoing in other normatively similar circumstances. To show what I mean, consider Adrian Dimow. Dimow was one of the subjects who disobeyed very early in Milgram's study. But Dimow reports that he didn't disobey out of concern for the learner, or because he thought it was the right thing to do. Instead, he disobeyed because he had a natural suspicion of authority figures as a result of being tracked by the FBI in his younger years, and he believed that the experimenters were trying to dupe him (Dimow 2004). Thus while he did the right thing in disobeying, he didn't do it for the right reasons, and so his actions aren't good evidence that he possessed sufficient commitment to V. It could still be the case that he would have harmed the learner if he thought the experimenters were genuine. And likewise, it could also be the case that other subjects

who successfully disobeyed might nevertheless obey some other authority figure in another setting, or cause harm when there was strong social pressure to do so. Disobeying in Milgram's original experiment is necessary but by no means sufficient for possessing the standing to blame subjects who obeyed.[13]

This is admittedly rather speculative, and I do not intend to argue the situationist experiments show that most people lack the standing to blame even if they do the right thing in difficult circumstances. But I think it is worth identifying that many of us are much less consistent and less committed to certain moral values than we like to think. We regularly overestimate our tendency to do the right thing for the right reasons, whether we would in fact do the right thing when the chips are down, and should be aware of this when criticising someone for failing to comply with morality's demands.

10. Conclusion

In this paper I examined the relationship between the situationist experiments, subjunctive hypocrisy and standing to blame. Using Todd's account of standing to blame, I argued that subjunctive wrongdoers lack the standing to blame if their subjunctive wrongdoing is attributable to a lack of commitment to the values which would condemn that wrong. I then argued that the fact that the obedient subjects in Milgram's experiments were distressed and had a history of not harming people is not sufficient to show that they were committed to the relevant values. The fact that these subjects had the capacity and opportunity to do the right thing for the right reasons, coupled with their severe wrongdoing, and low cost of disobeying shows they were not sufficiently committed to the relevant values. Since these subjects were ordinary citizens, and since Milgram's experiments have been replicated numerous times, this means that 80% of us lack the standing to blame these obedient subjects. I considered objections that this result is tantamount to excusing wrongdoers, promotes unethical behaviour, or is too revisionary. I argued in response that this doesn't feature more regularly in our moral practices only due to our epistemic limits and competing considerations, that we

[13]Doris (2002) acknowledges this point: 'Nevertheless, each subject was observed only in a single trial. Damn the obedients and hail the defiants if you will; the experiment does not motivate confidence about how particular subjects would behave in [other relevant trait-eliciting] situations. There's little reason for confidence that the disobedient subjects, however inspiring their behavior in the experiment could be counted on to exhibit Socratic self-mastery in other situations' (p. 49).

are still able to have a number of appropriate responses to wrongdoing, and that being less prone to blaming others could lead to a number of positive outcomes.

Acknowledgements:

Thanks to Luke Russell, Caroline West, Dana Nelkin, Isabelle Wentworth, Daphne Brandenburg, and Rosalind Chaplin for comments on earlier drafts and helpful discussion. Thanks also to audience members at the 2017 Australasian Association of Philosophy conference for their helpful questions.

Disclosure statement

No potential conflict of interest was reported by the author.

ORCID

Adam Piovarchy http://orcid.org/0000-0002-5169-2030

References

Badhwar, N. K. 2009. "The Milgram Experiments, Learned Helplessness, and Character Traits." *The Journal of Ethics* 13 (2-3): 257–289.

Blass, T. 1999a. "The Milgram Paradigm After 35 Years: Some Things We Now Know About Obedience to Authority." *Journal of Applied Social Psychology* 29 (5): 955–978. doi:10.1111/j.1559-1816.1999.tb00134.x.

Blass, T. 1999b. *Obedience to Authority: Current Perspectives on the Milgram Paradigm*. New York: Psychology Press.

Brink, D. O. 2013. "Situationism, Responsibility, and Fair Opportunity." *Social Philosophy and Policy* 30 (1): 121–149.

Brink, D. O., and D. K. Nelkin. 2013. "Fairness and the Architecture of Responsibility." In *Oxford Studies in Agency and Responsibility*, edited by D. Shoemaker, 284–313. New York: Oxford University Press.

Burger, J. M. 2009. "Replicating Milgram: Would People Still Obey Today?" *American Psychologist* 64 (1): 1–11. doi:10.1037/a0010932.

Chaplin, W. F., and L. R. Goldberg. 1984. "A Failure to Replicate the Bem and Allen Study of Individual Differences in Cross-Situational Consistency." *Journal of Personality and Social Psychology* 47 (5): 1074–1090.

Coates, J. D., and N. A. Tognazzini. 2013. "The Contours of Blame." In *Blame: Its Nature and Norms*, edited by J. D. Coates and N. A. Tognazzini, 3–26. New York: Oxford University Press.

Cohen, G. 2006. "Casting the First Stone: Who Can, and Who Can't, Condemn the Terrorists?" *Royal Institute of Philosophy Supplement* 81 (58): 113–136.

Darley, J. M., and C. D. Batson. 1973. ""From Jerusalem to Jericho": A Study of Situational and Dispositional Variables in Helping Behavior." *Journal of Personality and Social Psychology* 27 (1): 100–108.

Darley, J. M., and B. Latane. 1968. "Bystander Intervention in Emergencies: Diffusion of Responsibility." *Journal of Personality and Social Psychology* 8 (4): 377–383.

Darwall, S. L. 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, Massachusetts: Harvard University Press.

Dimow, J. 2004. Resisting Authority: A Personal Account of the Milgram Obedience Experiments. *Jewish Currents*. http://www.jewishcurrents.org/2004-jan-dimow.htm.

Doliński, D., T. Grzyb, M. Folwarczny, P. Grzybała, K. Krzyszycha, K. Martynowska, and J. Trojanowski. 2017. "Would you Deliver an Electric Shock in 2015? Obedience in the Experimental Paradigm Developed by Stanley Milgram in the 50 Years Following the Original Studies." *Social Psychological and Personality Science* 8 (9): 927–933.

Doris, J. M. 2002. *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press.

Edwards, D. M., P. Franks, D. Friedgood, G. Lobban, and H. Mackay. 1969. *An experiment on obedience*. Doctoral thesis, University of the Witwatersrand, Johannesburg, South Africa.

Friedman, M. 2013. "How to Blame People Responsibly." *Journal of Value Inquiry* 47 (3): 271–284.

Fritz, K. G., and D. Miller. 2018. "Hypocrisy and the Standing to Blame." *Pacific Philosophical Quarterly* 99 (1): 118–139.

Fritz, K. G., and D. J. Miller. 2019. "When Hypocrisy Undermines the Standing to Blame: a Response to Rossi." *Ethical Theory and Moral Practice* 22 (2): 379–384.

Gino, F., M. I. Norton, and D. Ariely. 2010. "The Counterfeit Self: The Deceptive Costs of Faking It." *Psychological Science* 21 (5): 712–720. doi:10.1177/0956797610366545.

Hartshorne, H., M. A. May, and J. B. Maller. 1929. *Studies in the Nature of Character*. New York: The MacMillan Company.

Isen, A. M., and P. F. Levin. 1972. "Effect of Feeling Good on Helping: Cookies and Kindness." *Journal of Personality and Social Psychology* 21 (3): 384–388.

Isserow, J., and C. Klein. 2017. "Hypocrisy and Moral Authority." *The Journal of Ethics and Social Philosophy* 12: 191–222.

King, M. 2015. "Manipulation Arguments and the Standing to Blame." *The Journal of Ethics and Social Philosophy* 9 (1): 1–21.

Latané, B., and J. M. Darley. 1970. *The Unresponsive Bystander: why Doesn't he Help?* New York: Appleton-Century Crofts.

Levy, N. 2014. *Consciousness and Moral Responsibility*. New York: Oxford University Press.

Levy, N. 2015. "Defending the Consciousness Thesis: a Response to Robichaud, SriPada, and Caruso." *Journal of Consciousness Studies* 2 (7-8): 61–76.

Macnamara, C. 2013. "Taking Demands Out of Blame." In *Blame: Its Nature and Norms*, edited by J. D. Coates and N. A. Tognazzini, 141–161. New York: Oxford University Press.

McKenna, M., and B. Warmke. 2017. "Does Situationism Threaten Free Will and Moral Responsibility?" *Journal of Moral Philosophy* 14 (6): 698–733.

Milgram, S. 1963. "Behavioral Study of Obedience." *The Journal of Abnormal and Social Psychology* 67 (4): 371–378. doi:10.1037/h0040525.
Milgram, S. 1974. *Obedience to Authority: an Experimental View*. London: Tavistock Press.
Miller, A. G. 1986. *The Obedience Experiments: A Case Study of Controversy in Social Science*. New York: Praeger Publishers.
Mischel, W., and P. K. Peake. 1982. "Beyond Déjà vu in the Search for Cross-Situational Consistency." *Psychological Review* 89 (6): 730–755.
Nagel, T. 1979. *Mortal Questions*. Cambridge: Cambridge University Press.
Nelkin, D. K. 2005. "Freedom, Responsibility and the Challenge of Situationism." *Midwest Studies in Philosophy* 29 (1): 181–206.
Nelkin, D. K. 2014. "Moral Luck." In *The Stanford Encyclopedia of Philosophy (Summer 2019 Edition)*, edited by E. N. Zalta. https://plato.stanford.edu/archives/sum2019/entries/moral-luck/.
Perry, G. 2013. *Behind the Shock Machine: The Untold Story of the Notorious Milgram Psychology Experiments*. London: Scribe Publications.
Piovarchy, A. J. 2020. Hypocrisy, Standing to Blame and Second-Person Authority (unpublished manuscript).
Rivera López, E. 2017. "The Fragility of our Moral Standing to Blame." *Ethical Perspectives* 24 (3): 333–361.
Roadevin, C. 2018. "Hypocritical Blame, Fairness, and Standing." *Metaphilosophy* 49 (1–2): 137–152.
Rossi, B. 2018. "The Commitment Account of Hypocrisy." *Ethical Theory and Moral Practice* 21 (3): 553–567.
Russell, L. 2009. "Is Situationism all bad News?" *Utilitas* 21 (4): 443–463.
Sabini, J., and M. Silver. 2005. "Lack of Character? Situationism Critiqued." *Ethics* 115 (3): 535–562.
Shanab, M. E., and K. A. Yahya. 1978. "A Cross-Cultural Study of Obedience." *Bulletin of the Psychonomic Society* 11 (4): 267–269. doi:10.3758/BF03336827.
Solomon, R. C. 2005. "What's Character got to do with it?" *Philosophy and Phenomenological Research* 71 (3): 648–655.
Sreenivasan, G. 2002. "Errors About Errors: Virtue Theory and Trait Attribution." *Mind; A Quarterly Review of Psychology and Philosophy* 111 (441): 47–68.
Talbert, M. 2009. "Situationism, Normative Competence, and Responsibility for Wartime Behavior." *Journal of Value Inquiry* 43 (3): 415–432.
Todd, P. 2019. "A Unified Account of the Moral Standing to Blame." *Noûs* 53 (2): 347–374.
Tosi, J., and B. Warmke. 2016. "Moral Grandstanding." *Philosophy & Public Affairs* 44 (3): 197–217.
Vargas, M. R. 2011. "Situationism and Moral Responsibility: Free Will in Fragments." In *Decomposing the Will*, edited by T. Vierkant, J. Kiverstein, and A. Clark, 325–349. Oxford University Press.
Wallace, R. J. 2010. "Hypocrisy, Moral Address, and the Equal Standing of Persons." *Philosophy and Public Affairs* 38 (4): 307–341.
Watson, G. 1996. "Two Faces of Responsibility." *Philosophical Topics* 24 (2): 227–248.