Taking Responsibility, Defensiveness, and the Blame Game

Pamela Hieronymi
December 3, 2021

Apparently this volume is like a jazz show, with some of us riffing on the themes of others. I am so pleased to be riffing on Paulina Sliwa's "Taking Responsibility." It is a rich, exciting essay, full of insight. Its subject is its title: taking responsibility. Its guiding question is, What is it to take responsibility for one's moral failures?

I would like to start by highlighting two gems from the opening page. Sliwa's first paragraph ends, "when we do fail, the right thing to do is to take responsibility." A very simple thought, but it bears reflection. Although her examples are all cases of moral failing, it seems to me the claim is unrestricted: if I fail in my artistic, athletic, or culinary endeavor, again, the right thing to do is to take responsibility for my failure—not to blame the paint, the equipment, or the produce. Not only is this the honest thing to do, but it is also the only way I will improve. (Of course, sometimes we are, in fact, foiled by the paint or the produce—honestly understanding the source of our failure is also important. I will return to this.)

Sliwa next states, and then argues against, an assumption found in much of the current philosophical literature on blame: "to take responsibility for a wrong is [simply] to blame oneself for it." "To blame" is understood, in this literature, to mean something like, "to hold responsible for wrongs done." Thus the literature seems to assume that *taking* responsibility for a wrong done is just the self-reflexive version of *holding* responsible—that to take responsibility is just to do, to yourself, what others would do to you in response to the wrongdoing. She calls this the "self-blame" account.

Once set out starkly and looked at from a bit of distance, the self-blame account seems, to me, bizarre—why would taking responsibility for one's *own* failure simply amount to responding to yourself as others would respond to you in light of that failure? Surely we should expect a self/other asymmetry, right here.

And, indeed, Sliwa provides a strong asymmetry in her own alternative account. She says, "I suggest that taking responsibility is a matter of owning the normative footprint of one's wrong… [to take responsibility is] to own [the] normative consequences." (1) Her idea is that moral failure "changes the normative landscape" by creating, e.g., duties to apologize, to repair, and even to feel badly—those duties constitute the "normative footprint." To take responsibility is "to recognize [these duties] and to commit to abide by them." (7)

(In my opinion, the word "normative" has become a vexing piece of jargon that usually adds only fog—it can typically be struck without any loss of comprehension and greater clarity can often be gained by replacing it with something more precise. I will either strike or replace it, when speaking in my own voice.)

According to Sliwa, then, whether or not one takes responsibility depends on how one relates to the "footprint" of one's own failure—to the various rights, obligations, and burdens that follow upon your failure. By starting with the footprint of moral failure, Sliwa's approach is similar to that of Julie Tannenbaum, who approaches moral responsibility by thinking about the need for various forms of repair, for responses to what she calls "moral inadequacy."[1] By focusing on the consequences of failure or inadequacy, Sliwa and Tannenbaum are able to navigate and explain aspects of moral responsibility that seem otherwise difficult or mysterious—in particular, they are able to say sensible things about cases of moral luck and cases of what Tannenbaum calls "mere moral failure."[2]

---

[1] Tannenbaum writes, "we should think of moral responsibility in terms of its connection to moral inadequacy: if one's action is morally inadequate (not merely not optimal), then one is morally responsible for that action." Julie Tannenbaum, "Moral Responsibility without Wrongdoing or Blame," in *Oxford Studies in Normative Ethics*, ed. Mark Timmons (2018), 126.

[2] There are also important differences between their approaches. Sliwa's topic is "taking responsibility" for *wrongdoing*. Tannenbaum is concerned with moral responsibility for harming or failing to aid, and she focuses especially on those cases in which there was no wrong, but instead either "mere failure" or moral luck. The two will disagree, I suspect, about whether certain failures count as wrongdoing.

WHY SELF-BLAME?

What made the self-blame account seem so appealing? Why would we have thought that taking responsibility is simply responding to yourself as others respond to you?

Note first that, if we adopt Sliwa and Tannenbaum's approach, it may seem that we are not responsible for our *successes,* nor for actions that are fully adequate.[3] Actions that have no particularly interesting interpersonal footprint, no distinctive impact on the network of interpersonal duties and expectations that run between us, will not be ones for which we can take responsibility, on Sliwa's account. In contrast, much of the current philosophical literature about moral responsibility thinks of an action for which we are responsible as an action such that it *could* have had such a footprint, whether or not it in fact does—so that even merely permissible, morally neutral actions (actions that do not impact others in any distinctive way) are still ones for which one is responsible, even though there may be nothing in particular for which one must *take* responsibility.

Why think of responsibility in this broader way, rather than focusing first on cases where there is something like liability or the need for repair? I suspect the difference in approach can be traced to the fact that many of us come to the topic of moral responsibility from concerns about free will and determinism. To explain:

Determinism suggests, to many people, that we lack a form of freedom required for moral responsibility—and it thereby suggests, to many people, that there we are not "responsible agents." A deterministic world can house what we could call "mundane" agents—thermostats, computers, other mechanistic movers and changers. Mundane agents are able to effect changes in the world in accordance with their own internal structure, but their activities are not apt targets of moral blame or punishment. If we are to be aptly blamed or justly punished, it seems, we must enjoy some further sort of agency, some further sort of control over our actions or some more robust ability to

---

[3] Though Tannenbaum gives a necessary, not sufficient, condition in the quotation in the note above.

do otherwise. This further form of agency, which we can call "responsible agency," is what seems to be incompatible with the truth of determinism.

Importantly, the thought is not that "responsible agency" is simply mundane agency with responsibility added; it is a different, more robust kind of agency, one that suits its bearer for blame and punishment.[4]

Note that if we *do* enjoy this more robust sort of agency, then it seems we will be exercising *that* form of agency—responsible agency—even when we act in morally neutral ways. Thus it can seem that we can be "responsible" even when doing nothing of consequence—because we are exercising our responsible agency.

Now: How would this focus on "responsible agency," so understood, lead us to the self-blame account? P. F. Strawson hoped to adjudicate a debate between those who thought that moral responsibility requires the falsity of determinism and some who did not.[5] In doing so, he introduced the extremely powerful idea of "reactive attitudes"—attitudes such as resentment, indignation, gratitude, or trust, which (unlike mere frustration, relief, or reliance) we have towards people whom we regard as responsible. It seems, to many people, that these attitudes mark responsible agency, in the sense just set out: to be a responsible agent is to be, generally speaking, someone towards whom such attitudes would be appropriate, should you do something of significance or consequence; and to hold someone responsible is to recognize their responsible agency by being ready to respond to them with these attitudes, if they should do something of

_____

[4] I argue against the thought that "reflective agency" will play this role in Pamela Hieronymi, "Reflection and Responsibility," *Philosophy and Public Affairs* 42, no. 1 (2014).

[5] Peter F. Strawson, "Freedom and Resentment," *Proceedings of the British Academy* xlviii (1962).

significance or consequence.[6] To blame someone is to respond with such attitudes. (In some cases, one might do more—one might impose some sanction or penalty or give some gift or prize. But those are special cases.)

From here it seems a very short step to think that to *take* responsibility is, most basically, to recognize your own responsible agency and that to recognize your own responsible agency is to stand ready to respond to yourself with the self-directed version of these attitudes: with guilt or remorse.

But, once we say this last part clearly and out loud, it seems (to me, at least) clearly mistaken. Sliwa deftly dispatches the self-blame account in part by pointing out the all-too-common phenomena in which someone rightly feels guilty and yet fails to take responsibility—in which their guilt instead motivates them to evade responsibility. Her example is John, who feels guilty for breaking his promise to attend a friend's event, but, when the friend brings it up, reacts defensively. John's reaction is fueled by guilt, and yet it is also a clear case of failing to take responsibility.

Sliwa and Tannenbaum instead approach the topic in a fresh way, without the background concern about freedom of the will, and so without the focus on the conditions for "responsible agency." They instead approach the topic with more purely moral or interpersonal concerns. In fact, it seems to me that "responsible agency" in the sense above does not play a role for either Sliwa or Tannenbaum. One might wonder: if we work backwards, so to speak, from the interpersonal footprint, will we ever encounter a need for the notion of "responsible agency"? If not, then we

_____

[6] Strawson has been interpreted as if he thinks that being the apt target of reactive attitudes is being a responsible agent. In fact, he is sometimes interpreted as claiming that the aptness of reactive attitudes constitute one as a responsible agent. This interpretation underlies the tendency to call his "type two" cases, in which the reactive attitudes are suspended altogether, cases of "exemption:" they are interpreted as cases in which a lack or or defect in agency has exempted one from responsibility.   I do not think this is the right way to understand Strawson—he is not concerned to identify "responsible agency" in the sense above; I believe he would deny that we need such a notion. I provide an interpretation of Strawson (and a reprint of his article) in Pamela Hieronymi, *Freedom, Resentment, and the Metaphysics of Morals* (Princeton: Princeton University Press, 2020).

should consider whether it is a philosophical invention created to address a philosophical problem in the pejorative sense—perhaps it is something we could do without.[7]

More immediately, neither Sliwa nor Tannenbaum is tempted to think of taking responsibility as standing ready to respond to one's own failure with self-blame or guilt. In fact, Tannenbaum explicitly says that blame would be inappropriate both in cases of what she calls "mere moral failure" and in cases of moral luck—even though she thinks that, in those cases, one is morally responsible. Again, Sliwa rejects the self-blame account with her example of John, who feels guilty but fails to take responsibility.

In considering John's reaction, Sliwa begins to address a topic that has loomed large in public life, lately, but that has been underrepresented in the philosophical literature: defensiveness. In my remaining comments, I am will focus on Sliwa's third section, "Making Excuses," in order to add some further thoughts on defensiveness, hopefully pushing the discussion a bit further. Considering her treatment of excuses will also help us to make contact—albeit only glancing—with legal practices.

OFFERING EXCUSES, BEING DEFENSIVE, AND TWO KINDS OF "NEGOTIATION"

Excuses, according to Sliwa, are facts that "bear on the size and nature" of the interpersonal footprint left by moral failure—in particular, an excuse "makes the size of the normative footprint smaller than it otherwise would have been." (10) For example,

> the fact that a wrongdoer… was provoked does not make her action any less of a wrong but it does make a difference to how we are entitled to feel about her wrongdoing, what kind of

---

[7] I suspect we can do without it. I suspect we should, instead, focus on a nearby idea, which I might call "answerable agency"—being such that your activities open you to questions and criticisms, where those questions and criticisms are answered or satisfied by appeal to the reasons for which you engaged in the activity. This is something beyond the mundane agency enjoyed by the thermostat. However, it is enjoyed by creatures who are *not* morally responsible, such as certain non-human animals or very young children (and so it is not "responsible agency"). More, not everything for which we are responsible is an instance of answerable agency. We are also responsible for states of affairs that fall within what I would call our "jurisdiction." I suspect that responsibility is layered on top of answerable agency, and answerable agency is compatible with the truth of determinism. See "Reflection and Responsibility."; *Minds That Matter* (in progress).

reparative actions she owes, and whether she is entitled to our good will, gossip, or dinner party invitation. (10–11)[8]

Sliwa claims, bluntly, that "making excuses is incompatible with taking responsibility." (10) I take it that her claim, here, concerns a certain activity, namely, *making* an excuse—that the mere *existence* of an excuse, the fact that a certain excuse exists and in some way mitigates one's responsibility, is compatible with taking responsibility for that which is not excused. Sliwa sometimes talks, instead, of "offering" an excuse, but it seems she treats "making" and "offering" as equivalent—both are the activity of presenting an excuse to another person.[9]

But: why should making an excuse, so understood, be incompatible with taking responsibility? Why isn't it enough to "own" the *actual* footprint, whatever its size turns out to be after the excuses are rendered? Why must one, in addition, refrain from presenting the facts that reveal its true size?

Sliwa says a few things that might answer this question. She says, first, that offering an excuse is "not just providing further information about our wrongdoing and our motivations for it," but is doing so "with the aim of mitigating the normative fallout of our wrong." (11) Thus, to offer an excuse is not just to present the excuse, but to do so with this particular aim.

But, why would *this* aim be problematic? Aiming to "mitigat[e] the normative fallout of our wrong" is not necessarily aiming, narrowly or selfishly, to reduce one's own burdens or improve one's own appearance—the aim need not be self-protective or defensive. "Mitigating the normative fallout of our wrong" would include such things as mitigating misunderstanding about the actual damage done ("Oh, wait—the vase didn't break; it's fine") or mitigating the hurt feelings caused by pointing out facts that reveal misunderstanding ("I'm so sorry—I didn't know about your

---

[8] Sliwa includes "acted under duress"—but it seems to me that duress does make the wrong less of a wrong, and sometimes no wrong at all. Throwing the cargo overboard would normally be impermissible, but, during the life-threatening storm, it is the right thing to do. (It also seems odd to think of oneself as *entitled* to gossip or dinner invitations.)

[9] For Sliwa's treatment of excuses, see Paulina Sliwa, "The Power of Excuses," *Philosophy and Public Affairs* 47, no. 1 (2019).

brother").[10] In mitigating the fallout in these ways, one need not be aiming either to deny one's own failing or to reduce one's obligations for self-interested or self-protective reasons.

Sliwa seems to suggest that, if you offer information in this non-self-protective spirit, you are offering what she calls an "explanation," not an excuse. As evidence, she points to the fact that, when you give the information, you may say, explicitly, "that's an explanation, not an excuse."

I agree that you might say this—in fact, saying it is often a good idea—but it seems to me that you will say it because you see the need to signal that you are still ready to take responsibility and that phrase accomplishes that end. Still, so long as the "explanation" you give does, in fact, "[make] the size of the normative footprint smaller than it otherwise would have been," then, working with the terminology as Sliwa presents it, the explanation is still *technically* an excuse, and you are offering it. Calling it an explanation, in your conversation, is simply a way of avoiding further misunderstanding.

However, we can, on Sliwa's behalf, isolate those offerings of excuse, like John's, that do have a self-protective or defensive motivation. Let us artificially call these cases of "making excuses," and use "offering excuses" to pick out the broader category of presenting information in order to mitigate the fallout of a wrong. It will then be true that making excuses, as an activity, and taking responsibility, as an activity, have opposing aims—because "making excuses" has now been defined as offering an excuse with the aim of protecting yourself from the burdens of taking responsibility.

But even this opposition of aims does not yet secure incompatibility. If (as Sliwa seems to assume) there is a fact of the matter about the size and shape of the footprint, then someone might make an excuse in order to protect themselves from having to take responsibility for anything

---

[10] Importantly, we should not expect that understanding our spotless intentions should eliminate all hurt feelings or wrong done. Often what matters is not whether your intentions were good but instead how your words or actions impacted the other—and that impact can stand as something for which you should take responsibility, despite your good intentions. Such cases are especially likely when the person impacted is subject to, e.g., systematic racism, homophobia, etc. Sliwa and Tannenbaum's approach is especially well suited to capture this fact. Still, we can grant all this, I think, and still leave room for the above sort of cases.

beyond the *true* footprint of their failure—in order to protect themselves from taking *too much* responsibility. In some cases, e.g. in abusive or unhealthy relationships, such self-protection may be appropriate—it may be important to avoid being over-responsible. In other cases, it may simply be establishing the important facts: I shot the sheriff, but I did not shoot the deputy.[11] In yet other cases, someone who self-protectively insists on identifying the footprint very precisely might show an ungenerous or stingy attitude—but even so, they are not, it seems to me, failing to own the footprint of their wrongdoing. They are simply refusing to own anything more. Even this stingy, defensive posture does not seem to me strictly incompatible with what Sliwa identifies as taking responsibility: owning the footprint of one's wrongdoing.

What *would* be incompatible with owning the footprint would be what we might call "making *up* excuses"—presenting considerations in order to convince the other person that the footprint is smaller than it appears to them, for self-protective reasons, while nonetheless in some way believing or suspecting that it is in fact larger. This, it seems, is what John is doing.[12] Alternatively, someone might make up excuses without having any beliefs about the size the of footprint: they might simply want to reduce its appearance as much as possible, without any concern about its true size. Making up excuses is incompatible with taking responsibility because the person who makes up excuses is not interested in the footprint, itself, but rather attempting to own as little as they think they can get away with owning. This is, indeed, incompatible with a willingness to take responsibility as Sliwa has defined it.[13]

---

[11] Credit for the Bob Marley reference goes to Mark C. Johnson.

[12] John feels guilty, and so he in some sense believes it is as large as it seems to his friend, but is trying to ignore that or to convince both the other person and himself that he has done no wrong.

[13] Note: Making up excuses is incompatible with taking responsibility, *even though* the one who makes up excuses might be willing to take responsibility for whatever they *cannot* get away with. It is incompatible with taking responsibility because the person is willing to own *only* whatever they cannot avoid by making up their excuses.

I would now like to focus on a complication that Sliwa does not consider—one that will connect to some of the legal issues in this volume. In the aftermath of a wrong there will likely be disagreement about the size and shape of its footprint. The footprint, as you see it, may not be the footprint as it appears to others. Until there is some rough agreement about what happened and how it affected various things of significance (or else clarity that no such agreement will be reached), it may not be clear exactly which obligations, etc., are owed. In fact, there may not yet be a fact of the matter about the footprint: the exact size, shape, and content of the footprint may await the process of figuring out how the events are seen from each side, how important the damages are to each one, whether or how learning about the other point of view changes one's own sense of the events and their importance, and whether, or to what extent, agreement about such things can be reached. In fact, even after agreement is either reached or forgone, the footprint may have unavoidably fuzzy edges.[14]

Some of what Sliwa says seems in tension with the thought that the footprint may need to be made determinate though conversation—in fact, some of what she says seems to suggest that engaging in such a conversation would, by itself, be a failure to take responsibility. She says,

> the person taking responsibility sees the question of which normative changes have taken place as settled and commits to acting in light of them. In contrast, the person who engages in negotiations about the normative footprint must regard this question as open, as something that can be contested. Taking responsibility and offering excuses are opposed to one another because they involve two incompatible attitudes towards the normative footprint. (12)

Sliwa seems to think that, to "negotiate" the size of the footprint is to "contest" it. In fact, she earlier seems to identify negotiating the footprint with failing to take responsibility:

> [when we make excuses] we [do] so with the aim of mitigating the normative fallout of our wrong. In other words, to [make] excuses is not merely to *assert* something but to *negotiate*….

---

[14] Sliwa may be thinking that clarifying the contours of the footprint is, itself, part of the footprint, so to speak—that the first task in addressing the fallout is to identify the fallout. If so, then it seems that a wrong will create a footprint for the person wronged—that they may have obligations and responsibilities in clarifying the fallout. Sliwa will then have the task of clarifying the difference between the victim and, say, the Dean. It seems sensible to think that the Dean (or other third parties) can take responsibility for the fallout of a wrong. It seems odd to think that the victim is required to do the same.

we are not merely trying to change the other party's mind about our motivation or circumstances of our wrongdoing but about its normative consequences: what kind of apology we owe, whether compensation is required, what kind of attitude others are entitled to take towards us. We are negotiating about the size of the normative footprint. (11)

Sliwa seems to think that to negotiate is to do something bad, something like making up excuses —attempting to reduce the footprint as much as possible. As I hear it, though, "negotiate" can describe two very different possibilities (as well as range of cases between them): Negotiations can be cooperative activities or they can be low-level power struggles ("contests").[15] In the latter case, they might be self-protective or even sophistical attempts to evade responsibility, but in the former, it seems to me, they need not be—even if what is being negotiated is the size and shape of the footprint of one's wrongdoing.

Let us step back for a moment to think about negotiations, generally, starting with the cooperative sort. Negotiations are cooperative when each party is, with good will, seeking an outcome that is acceptable to each of the others. Each one is not only willing to be honest and upfront about their own needs, interests, and desires but also willing to take the needs, interests, and desires of others with equal seriousness.[16] Cooperative negotiations require trust and respect—perhaps a daunting degree of both (especially if the stakes are higher than, say, how to spend a free evening). But, importantly, all that trust, respect, and general good will does not preclude negotiation—it does not eliminate the need to put forward your own interests, etc., honestly and without either minimizing them or enlarging them, understanding that others will put forward theirs, honestly and without either minimizing or enlarging them, in the process of finding the mutually acceptable outcome. Good will does not preclude negotiation because people are idiosyncratic—with different

---

[15] "Low-level" power struggles because conducted in speech rather than outright blows (high-level power struggles), in behind-the-back conspiracy (subterranean power struggles), or in a brute refusal to speak or to act (passive-aggressive power struggles). (Cf. Clausewitz' adage about war as diplomacy by other means; it might be put the other way around.)

[16] This is, roughly, the basic form of contractualist reasoning as understood by Scanlon in T. M. Scanlon, *What We Owe to Each Other* (Cambridge, MA: Harvard University Press, 1998). One result of taking all interests with equal seriousness may be the establishment of general prerogatives (for example, for individuals to favor their near and dear).

preferences, interests, needs, etc.—and until we have all that information on the table, so to speak, we will be stuck with some pre-fab, one-size-fits-all outcome that may be worse than what we could work out between us by negotiating. To illustrate: before talking, the right thing for mutually caring, deferential people to do might be to split each of the tasks equally between us—that is the pre-fab solution. But, if I prefer the first and you prefer the second, we can do better. There is no need to settle for the one-size-fits-all outcome. In fact, in such exchanges we learn about one another; we are drawn into partnership, sometimes even intimacy.[17]

It seems to me that one could *cooperatively* negotiate about the size and shape of the interpersonal footprint left by one's own wrongdoing, while at the same time standing ready to take responsibility for it once its size and shape have been clarified. Again, you might simply be wanting to ensure that all the facts are known by the other (the vase didn't break; I didn't know about your brother; and I definitely did not shoot the deputy), or you may be wanting to ensure that you understand how your actions have affected the other, how they see things, how they react to your way of seeing things, whether certain facts will make a difference to them or not, etc. Though you are negotiating, in a sense, you are not making up excuses—you are, instead, seeking to discover, and perhaps to make determinate, the size and shape of the footprint, in order to take responsibility for it.

Alternatively, and at the other extreme, negotiations can be low-level power struggles in which at least one side seeks their preferred outcome, unconcerned about reaching an agreement that is acceptable to others (except insofar as such acceptance is required by some "leverage" held by the

---

[17] The one-size-fits-all framework is, roughly, contractualist morality in the narrow sense—the outcome of such a negotiation, hypothetically conducted, based on personal and "generic" reasons. Actual, cooperative negotiation can do one better. In fact, in a cooperative negotiation individuals might settle on an outcome that is not, strictly, fair: it may be that one party is willing to do to more than their fair share, given the situation. An honest negotiation will allow for their generosity or accommodation to be acknowledged and appreciated. (It is this possibility the led Scanlon to formulate his contractualism in terms of principles that "no one could reasonably reject" rather than principles that "could be reasonably accepted"—he did not want to call the generous unreasonable. Note, too: this complication reveals the inadequacy of the word "normative:" the "normative footprint" need not follow the usual "normative" demands. We might do better to say that the "interpersonal footprint" need not follow the usual "moral demands"?) Of course, stepping away from fairness is fraught business (cf. the usual gendered division of household labor), and may instead hamper intimacy.

other party).[18] Given the fixation on one's own preferred outcome and the lack of concern for mutual acceptability, a deficit of honesty is likely: certain facts may not be volunteered, irrelevancies may be brought up and put forward to distract, stories may be spun. Those gifted at what Harry Frankfurt calls "bullshitting" will be particularly well placed in these contests—because they are unconcerned with truth, they are not hindered by it when speaking.[19] In a power-struggle-style negotiation, at least one of the parties aims to win the argument rather than to discover the truth or to arrive at a mutually acceptable outcome.[20]

What is it to "win the argument," in the sense at issue here? One will "win the argument," in this sense, when the other party yields—when the conversation ends with your position established as the "common ground" in the technical sense, as that which will be *treated as true*, going forward.[21] Importantly, winning the argument, in this sense, does not require actually convincing the other party. Nor does it require discovering, changing, or making determinate the *actual* footprint. It only requires that we all leave the conversation with the winner's preferred interpretation of that footprint as the established common ground.[22]

---

[18] There is, of course, a range of cases between these extremes, and the cooperative sort of negotiation can easily degenerate into the power-struggle sort. People are frequently *somewhat* concerned to find an outcome acceptable to the other, but are willing to take others into account only within certain limits. (One might think of those limits as a kind of leverage had by the other party.)

[19] Nor are they likely to recognize their own disinterest in reaching mutually acceptable outcomes. See Harry Frankfurt, "On Bullshit," in *The Importance of What We Care About* (Cambridge University Press, 1988).

[20] Notice that the power struggle need not be driven by *selfishness*: it can be conducted by someone who genuinely seeks to establish, not just what they want or what would be good for them, but what they sincerely believe to be best overall. Someone who is both confident they know what is best and also determined to implement it regardless of others' differing opinions will engage in a conversation only for the purpose of (avoiding the appearance of unilateral action by) gaining the other's assent. In such conversations, forthrightness and honesty will be significant handicaps, and therefore frequently foregone. (The difficulty here is faced both by well-intentioned salespeople and by confident and caring doctors.)

[21] For the notion of "common ground" see Robert Stalnaker, "Assertion," in *Pragmatics*, ed. Peter Cole (New York: New York Academic Press, 1978; reprint, Robert Stalnaker, Context and Content OUP 1999). and subsequent literature.

[22] Those who like to win arguments often assume that, when they have done so, they have also convinced the other party —or that, if have not, that is due to the vice or the stupidity of the other party. Often enough, neither is true.

I agree with Sliwa that *this* kind of negotiation about the footprint would be incompatible with taking responsibility—those engaged in it are not concerned with owning either the true footprint or the one made determinate though cooperative interaction, but rather with owning whatever they can establish by winning the argument. They have the same aim as those who make up excuses.[23]

We can note that the adversarial nature of our legal system seems to ensure that legal disputes are negotiations of the power-struggle sort—each side attempts to win an argument. Does this mean that taking legal action, or defending oneself against a suit, is incompatible with taking responsibility for one's failures? Not necessarily—but it would be interesting to explore why or under what conditions it is or is not. Legal recourse is often the recourse of last resort. In addition, the institutionalization of the adversarial procedure, together with the employment of a third party (the judge) to establish the final "common ground," might change our sense of what does or does not count as either the footprint or owning it. I will not explore these questions further, but I hope what I say is suggestive of possible questions and possible answers.

---

[23] It seems to me that the power-struggle-style conversation about the footprint can come in three forms, two of which we have already considered: You may be making up excuses while in some sense believing the footprint is larger than you are presenting it, like John. Or, like Frankfurt's bullshitter, you may be making up excuses without any opinion about or concern for the actual footprint, but aiming only to establish, as common ground, the smallest possible interpretation of it. And, thirdly, like the salesperson or doctor of the earlier footnote, you may be concerned about the footprint but also dogmatically convinced of your own view of it. You might then enter the conversation only to ensure that your own interpretation is established as the common ground. This, too, seems to me a way of failing to take responsibility.

Why would this last case be failing to take responsibility? If there is a fact of the matter about the size, shape, and content of the footprint, and if you are correct about it, why, in that case, should dogmatically seeking to establish the correct footprint as the common ground be incompatible with owning it? So long as you are not *dishonest* in your dogmatic attempts at persuasion, why would this be a failure to take responsibility? I want to grant that you sometimes should uphold one's your view of the footprint (you did *not* shoot the deputy!), and that you sometimes show should do even if you cannot articulate exactly why (you cannot say exactly why you do not need to kiss him and make up, but you are sure you do *not*). Still, being dogmatically unwilling to reconsider one's view, from the beginning, does seem to me incompatible with taking responsibility—at least often. This might be evidence that the exact size and shape of the footprint is not given in advance, or it might be evidence that part of the footprint is engaging in some process of discover about the footprint.

THE BLAME GAME AND THE NAMELESS VIRTUE

I would like now to widen focus while continuing to consider defensiveness. A power-struggle style conversation about the size, shape, and content of the footprint is often only one manifestation of a larger power struggle that I will call "the blame game."

My go-to example of the blame game is the lecture course going badly—when students are not engaged and not learning. (I have some experience with this, unfortunately.) In such cases, blame typically abounds. The instructor is likely to blame the students. The instructor will think the students cannot expect to learn or to find the lectures engaging if they keep looking at their phones or surfing the web, or if they come to class without having done the reading, or if they do not ask questions to reveal their confusions. Meanwhile, the students are likely to blame the instructor. They will think that the instructor needs to make the material more accessible or more relevant to their interests, or to deliver the lecture in less of a monotone, or to prepare better visual aids. In blaming the other, each side is avoiding the burdens of improving the situation by changing themselves. The blame-game need not involve any conversation, but it is still a power struggle.[24]

Note that the "blame," in this case, is not the "blame" of the current philosophical literature. In the current literature, "blame" has become a technical term; it refers to responses to moral failure, with the "reactive attitudes" of resentment and moral indignation serving as central cases (as noted above). In contrast, the blame assigned by the instructor or the students need not be moral at all: neither side need think the other is failing *morally* (rather than professionally or academically). Rather, the "blame," in this case, is the "blame" of ordinary English: there is some kind of mess or problem that was the foreseeable result of someone's failure; that mess or problem is that person's "fault," they are "to blame" for it; and the mess or problem is also, therefore, presumptively, theirs

---

[24] It may be a subterranean one or a passive-aggressive one (see note 17).

to clean or to fix (cf., "The Dean is to blame for the delay").[25] In the blame game, each side assigns fault to the other, believing that whomever is at fault will also need to fix the problem—and thus, in the blame game, each side is not only failing to take responsibility themselves but also trying to shift the burdens of fault and fixing onto the other party.[26]

The blame game is something we would like to avoid. But, as noted, cooperative negotiations require a high degree of trust, honesty, and good will—in the wake of failure, those are often in short supply. In the wake of a failure, when the exact size, shape, and content of the footprint is not yet clear, *must* you attempt a conversation about it? Must you attempt to initiate a cooperative negotiation of the footprint? It seems to me not. Often enough, cooperative negotiation will be obviously beyond reach. Even if not entirely out of reach, you may find it a daunting and unattractive prospect, and it may be permissible to opt against. One might instead decide, simply and perhaps generously, to own a broad interpretation of the footprint. Often enough, doing so is appropriate, perhaps even praiseworthy.[27]

I believe that taking this option—generously owning a broad interpretation of the footprint—is what Susan Wolf has in mind with her "unnamed virtue," discussed by Sliwa at the end of her chapter.[28] Wolf seems to have in mind the opposite of the stingy person we encountered earlier, who is concerned to identify the precise size of the footprint and to own no more than that. The

---

[25] Sliwa's idea of a "footprint" fits nicely into an account of the blame game (as noted, her analysis seems easily extends to non-moral cases). When some mess or difficulty is your fault, you will typically incur obligations to fix it or to clean it up—typically, a mess creates a footprint to be owned, presumptively, by the one at fault for the mess. But not always. Cf. Sliwa's case of the Dean taking responsibility for the behavior of the faculty member.

[26] The above-mentioned "passive-aggressive power struggle" shifts burdens to the other party in a much more efficient way: you simply refuse to take them up. Why, then, are people drawn into conversation? I suspect it is because those skilled at argument will leave it feeling justified—not only their own eyes, but also in what they think should be eyes of the other. The passive-aggressive strategy is the one left to those who are less successful in winning the argument or who have less patience for arguments conducted in bad faith.

[27] One's decision may later turn out to be the first step in a negotiation—either a cooperative one or a power-struggle-style one.

[28] See Susan Wolf, "The Moral of Moral Luck," *Philosophic Exchange* 31, no. 1 (2001).

person with Wolf's virtue is happy to take on a bit more, is not terribly concerned about questions of fault or failure, except, perhaps, as it will help us understand how to improve or to fix (as when we are foiled by the produce or the paint).

Of course, as with any virtue, there is the possibility of missing the mark—you might be over-responsible, overly inclined to pick up burdens, thus over-burdening yourself and "enabling" others, leaving yourself open to being taken advantage of. Or, you might be too quick to dismiss questions of fault; you might take up a generous interpretation of the footprint as a way of deflecting those questions—you may prefer to take up obligations and tasks rather than to face your own failure. (You will then be unable to learn from those failures.) So it does seem that there is a virtue here— something that is to be done at the right time, to the right degree, in the right way, etc.[29]

Sliwa suggests a name for Wolf's nameless virtue: "being responsible." I am not convinced this is an apt name, for two reasons.

First, the person who misses the mean of this virtue, either by being stingy, by being over-responsible, or by wanting to avoid focusing on failures, is not failing to take responsibility. They are going wrong in a somewhat different way.

Second, "being responsible" will be a good name for Wolf's virtue only if we can distinguish it from something else we might naturally call "being responsible:" being conscientious, reliable, trustworthy, someone who takes their obligations seriously, someone to whom you would give a complex, difficult job. Often enough, those who are responsible, in this latter sense, *also* display Wolf's virtue. But not always. Sometimes those who are conscientious, etc., are so because they are very hard on themselves, and, because they are very hard on themselves, they are also very touchy about assignments of fault. (Sometimes the first person to whom you would want to give the job is the last person you would want to confront about their failure.) Sometimes when we say someone is

---

[29] Aristotle thinks of virtues as a "mean" between two extremes. Without suggesting that there is a single gradient with a "middle," I would like to put more emphasis than either Sliwa or Wolf on the two hazards just mentioned.

a "responsible" sort of person, we do have in mind Wolf's virtue—or, at least, her virtue is part of what we have in mind. But I agree with Wolf that it does not, quite, have its own name.

I close with a perhaps unexpected observation: We are now in a position to understand an otherwise puzzling disagreement among Amazon reviewers of a children's book, *The Berenstain Bears and the Blame Game*.[30] Reviewers include both many who praise the book for teaching children to take responsibility and many who criticize it for failing to do so. How could one story earn both reviews?

In the story, playing children break a vase and then break a window. After each debacle, the blame game begins: each child claims the other is at fault. A parent calls off the game, saying, "There is more than enough blame to go around." The parent then sets the children to clean up, while helping them with it. The children, it seems, are taught to set aside the blame game, and thereby to set aside the questions of fault, and instead to focus on fixing.

Critics see the story as failing to teach the children to take responsibility: the children are never made to articulate that for which they were individually at fault, and so they do not take up the tasks as part of a footprint that they have identified as the impact of their own failure. It seems, to the critics, that individual responsibility has been avoided.

Fans seem, instead, to see the story as teaching Wolf's virtue: Rather than sort out questions of fault in a fine-grained way, the children are taught to each accept some admittedly unspecified degree of it and simply take up the tasks contained in footprint.

Though the disagreement seems to me not only intelligible but also sensible, I side with the fans.

CONCLUSION

What have we learned? First, Sliwa's idea of taking responsibility as "owning the footprint" is an extremely fruitful, powerful one. Second, we might want to reconsider whether thinking about

---

[30] Stan Berenstain and Jan Berenstain, *The Berenstain Bears and the Blame Game* (New York: Random House, 1997). Thanks to Jenna Donohue for drawing my attention to this book.

"responsible agency" is the best way to think about responsibility. Third, we have seen that there is a question about whether the "footprint" of a wrong is given in advance or instead whether it must be made determinate through interaction, perhaps through conversation. Fourth, we have distinguished between two different kinds of conversation: a cooperative negotiation and a low-level power struggle. The power-struggle-style conversation may be part of a larger struggle I have called "the blame game" (where the "blame" in question is not that referred to in the current philosophical literature). Finally, we have noted that one may sometimes opt out of the blame game simply by owning a generous interpretation of the footprint. Doing so is sometimes, but not always, virtuous.

I have suggested that engaging in the power-struggle-style negotiation is generally incompatible with taking responsibility and also that legal disputes are power-struggle-style negotiations. However, I do not think it follows, straightforwardly, that being a party to a legal dispute is incompatible with taking responsibility. Rather, I noted in passing that the fact that legal recourse is a last resort, the institutionalized nature of the legal dispute, and the presence of a third party in the role of judge, may allow us to say that being party to a legal dispute is compatible with taking responsibility—but I have not further explored this matter. I leave it as an exercise for the reader.[31]

BIBLIOGRAPHY

Berenstain, Stan, and Jan Berenstain. *The Berenstain Bears and the Blame Game*. New York: Random House, 1997.

Frankfurt, Harry. "On Bullshit." In *The Importance of What We Care About*, 117–33: Cambridge University Press, 1988.

Hieronymi, Pamela. *Freedom, Resentment, and the Metaphysics of Morals*. Princeton: Princeton University Press, 2020.

———. *Minds That Matter*. in progress.

———. "Reflection and Responsibility." *Philosophy and Public Affairs* 42, no. 1 (2014): 3–41.

Scanlon, T. M. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press, 1998.

Sliwa, Paulina. "The Power of Excuses." *Philosophy and Public Affairs* 47, no. 1 (2019): 37–71.

Stalnaker, Robert. "Assertion." In *Pragmatics*, edited by Peter Cole, 315–32. New York: New York Academic Press, 1978. Reprint, Robert Stalnaker, Context and Content OUP 1999.

Strawson, Peter F. "Freedom and Resentment." *Proceedings of the British Academy* xlviii (1962): 1-25.

Tannenbaum, Julie. "Moral Responsibility without Wrongdoing or Blame." In *Oxford Studies in Normative Ethics*, edited by Mark Timmons, 124–28, 2018.

Wolf, Susan. "The Moral of Moral Luck." *Philosophic Exchange* 31, no. 1 (2001): 5–19.