

May 2018

Reactive Attitudes & the Value of Responsibility

Andrew Lichter

University of Wisconsin-Milwaukee

Follow this and additional works at: <https://dc.uwm.edu/etd>



Part of the [Ethics and Political Philosophy Commons](#)

Recommended Citation

Lichter, Andrew, "Reactive Attitudes & the Value of Responsibility" (2018). *Theses and Dissertations*. 1860.
<https://dc.uwm.edu/etd/1860>

This Thesis is brought to you for free and open access by UWM Digital Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UWM Digital Commons. For more information, please contact open-access@uwm.edu.

REACTIVE ATTITUDES & THE VALUE OF RESPONSIBILITY

by

Andrew Lichter

A Thesis Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Master of Arts

in Philosophy

at

The University of Wisconsin-Milwaukee

May 2018

ABSTRACT

REACTIVE ATTITUDES & THE VALUE OF RESPONSIBILITY

by

Andrew Lichter

The University of Wisconsin-Milwaukee, 2018
Under the Supervision of Professor Andrea Westlund

This paper argues against the family of “reactive accounts” of moral responsibility. On such accounts, which take their cue from P.F. Strawson’s influential “Freedom and Resentment,” *being* morally responsible is properly understood in terms of being *held* responsible, which in turn is properly understood in terms of a set of moral emotions and their associated practices. This way of understanding responsibility re-frames apparently metaphysical questions about whether we are responsible in normative terms, as questions about whether and why these practices are permissible or required. I argue that we are responsible because we in some sense affirm the value of being held responsible, but that the reactive account fails to accommodate this fact about our responsibility. Specifically, this is the case because the moral emotions at issue do not seem to have a necessary connection to the way of regarding other people that must be implicit in our responsibility practices if we are to affirm their value. For one thing, there is no necessary connection between certain commitments and certain affective dispositions. For another, it is not hard to come up with instances in which these attitudes betray an obvious failure to share in the commitments that make responsibility practices valuable. Finally, I argue for the point that such a necessary connection between evaluative commitments and responsibility practices is crucial to a theory of the latter insofar as we want the theory itself to convey these commitments.

To my parents,
whose unwavering support
both humbles and motivates me.

TABLE OF CONTENTS

INTRODUCTION	1
§1 THE VALUE OF RESPONSIBILITY.....	3
1.1 <i>Strawson & the Justification for Responsibility Practices</i>	3
1.2 <i>The Value of Responsibility</i>	7
1.3 <i>Relationships & Responsibility</i>	8
1.4 <i>Commitment, Demands and Trust</i>	9
1.5 <i>Relational Morality and the “Problem of the Stranger”</i>	11
§2 AFFECT & ENDORSEMENT.....	16
2.1 <i>Endorsing Commitments vs. Endorsing Emotions</i>	16
2.2 <i>Why the Reactive Account Needs Reflective Endorsement</i>	17
§3 CORRUPTIONS OF RESENTMENT.....	19
3.1 <i>Interpreting Reactive Attitudes</i>	19
3.2 <i>Corruptions of Resentment</i>	21
§4 RESPONSIBILITY & MODALITY	25
4.1 <i>Vulcans, Cheerleaders & Valuing</i>	25
4.2 <i>Theories of Better and Better Theories</i>	29
CONCLUSION.....	32
REFERENCES.....	34

INTRODUCTION

That emotions play a crucial role in our moral life is undeniable. The nature of this role, however, is difficult to identify. And determining their *proper* role is more difficult still. A developing tradition gives the moral sentiments pride of place in moral life, holding that foundational moral concepts such as responsibility and blameworthiness are appropriately explained and understood in terms of their connection to attitudes such as guilt, resentment and indignation.¹ In this paper, I deny that such sentiments should have this pride of place in a theory of responsibility. By considering the nature of the commitments of which responsibility practices are a part, I argue that the reactive attitudes are not a necessary component of what is justified by our higher-order endorsement of these practices. As such, it is misguided to think of these attitudes as a fundamental element of moral and interpersonal responsibility. More basic is our commitment to reciprocal relationships, which both gives rise to and justifies the practice of holding responsible. The reactive attitudes bear no necessary connection to this commitment and as such are not an essential part of what we endorse when we endorse responsibility practices.

I will proceed as follows. First, I will frame the inquiry at hand in terms of P.F. Strawson's claim that we would not forego the entanglements of interpersonal commitment even if we could. This suggests that responsibility practices are justified by way of reflective endorsement: we have a higher-order commitment to an investment in a certain way of relating to one another. In §1, I consider an example to bring out why we regard holding and being held responsible to be valuable. By examining why we think a framework for responsibility is justified in cases where being held responsible is unpleasant, I argue that we think our responsibility practices are valuable because they

¹ This tradition is rooted in Strawson (1982). Watson (2004b) is a crucial development of the target view. I will primarily concern myself with Wallace (1994), McKenna (2012) and McKenna (2017). All of these theorists see themselves as improving on Strawson's original view. As such I will usually lump them under the umbrella of "Strawsonian accounts" or use Wallace's preferred "reactive accounts." I also discuss Darwall (2006), who endorses and defends some Strawsonian machinery.

are implicated in a commitment to conducting relationships defined by reciprocity. In §2, I suggest that endorsing this commitment entails nothing with respect to an endorsement of being subject to the reactive attitudes. In §3, I argue that the picture of the reactive attitudes upon which the reactive account relies is overly sanitized. In some cases, these attitudes do not involve regard for others, rooted in our own commitment to reciprocity, in terms of their capacity to conduct reciprocal relationships. As such, they are not fit to play the central role in a conception of responsibility that many philosophers afford them. In §4, I consider an objection to the demand for a necessary connection between the commitments we endorse and the practices to which they supposedly give rise and thereby justify. Here, I argue that the reasons we have for endorsing responsibility mean we also have reason for insisting on a conception of it that is characterized by such a necessary connection.

I am not the first to argue against the Strawsonian view that moral sentiments are the keystone in a conception of responsibility.² My aim in this paper is to argue for this familiar conclusion in a new way by interrogating the commitments that lead us to think of being and holding one another responsible as indispensable to common life. This approach may seem to risk illegitimately blurring the distinction between a description of responsibility practices and a justification of them. But I see such an elision as appropriate in this domain. It is in virtue of what matters to us that we hold others responsible: if we were not committed to the value of relating to one another in terms of reciprocity and mutual regard, we would make no demands on others' conduct. As such, an account of our responsibility practices must accommodate an essential tie to the commitments that give rise to these practices.

² I am most notably in the company of Scanlon (1998, 2008). See also Arpaly (2006), Gaus (2011), Korsgaard (1996), Pereboom (2001).

1.1 | *Strawson & the Justification for Responsibility Practices*

At the center of the contemporary literature on moral responsibility is P.F. Strawson's "Freedom and Resentment." The Strawsonian account of moral responsibility centers the reactive attitudes—paradigmatically resentment, indignation and guilt—and argues that responsibility can be analyzed in terms of them. Michael McKenna summarizes the two crucial Strawsonian theses as follows:

- (1) Being morally responsible must be understood by reference to the nature of holding morally responsible; and
- (2) Holding morally responsible ought to be understood by reference to a particular range of moral emotions and their related practices.³

My concern is primarily with (2), which tells us that to experience and act on the moral emotions is to hold responsible; to be their object is to be responsible.⁴ Although in Strawson's long shadow the link between holding others responsible and our sentiments may seem nigh unto a given, I think the claim that one's responsibility is appropriately understood in terms of others' sentiments should strike us as *prima facie* unsettling. This is true even once we outfit the sentiments in question with all of the complex machinery proposed by reactive account theorists such as McKenna and R. Jay Wallace, to whom it is crucial that the reactive attitudes are characterized by a connection to propositions about the character and actions of their objects. At least we may wonder what is doing the work: the emotions in question, or all of the machinery to which we are told they are attached.

One way to approach this question is by asking about the terms on which our responsibility practices are justified. Whether or not they are sympathetic to Strawson's views, philosophers often hone in on his claim that the reactive attitudes need not be justified insofar as they are part of a

³ McKenna (2012, 31). The phrasing is essentially McKenna's.

⁴ Strawson (1982, 78)

“general framework of human life” which is “not something that can come up for review.”⁵ But Strawson also makes a further point. Even if we *could* bring our participant attitudes up for review (again, a possibility he denies), he says, we would re-up. Upon reflection, we find that our lives are richer in virtue of an investment in interpersonal entanglements. As such, we would never opt to forego these entanglements.

Wallace calls this Strawson’s “pragmatic argument,” so named because it is framed as an answer to the question of what would be *rational* to do in the face of determinism.⁶ On this front, Strawson tells us that “we could choose rationally only in the light of an assessment of the gains or losses to human life, its enrichment or impoverishment; and the truth or falsity of determinism would not bear on the rationality of *this* choice.”⁷ Wallace interprets this as a sort of impersonal query about the usefulness of responsibility practices as a whole. On his reading of Strawson, the pragmatic argument explicitly looks to operate outside of the moral realm because the question to which it provides an answer is posed from *outside* the framework of moral standards that we deploy when we actually go about holding one another responsible. The question seeks nonmoral justification for our responsibility practices which in turn would not be undermined by the truth of determinism.

Wallace argues that this way of justifying our practices does not yet secure the right result given the dialectic. On his view, the incompatibilist charge is levied from *within* the framework of normative ethics. The fundamental moral principle at play is fairness. The concern of the incompatibilist is that, just as we think it would be unfair to hold someone responsible who was insane or a child, the truth of determinism would render it unfair to hold anyone responsible at all insofar as her actions and character are appropriately attributable to her heredity and environmental.

⁵ Strawson (1982, 70). For an insightful criticism of this point, see Russell (1992).

⁶ Wallace (1994, 99-100)

⁷ Strawson (1982, 70)

If our responsibility practices are to be justified to the incompatibilist, it must therefore be on moral grounds, for the incompatibilist charge (or at least the most reasonable version of it) is a moral one. The pragmatic argument, meanwhile, serves merely to establish that responsibility practices enrich human life, but not yet that determinism would not render them unfair regardless of the purported enrichment.⁸

Now, even on this reading, the pragmatic argument nonetheless strikes me as a greater blow to incompatibilism than Strawson's merely naturalistic claim that we are *unable* to give up our responsibility practices. Surely, the argument that the practices in question serve to enrich our lives has greater justificatory force than the claim that we couldn't change our ways even if we wanted to do so. Beyond this point though, it is not clear to me why Wallace insists on understanding concerns about the enrichment or impoverishment of human life as purely pragmatic concerns. One explanation is that he is disposed to understand moral concerns primarily on deontological terms, and that talk of enrichment may strike us as the domain of human flourishing. For the professedly Kantian Wallace, that responsibility practices may make things go better for us all things considered does not yet mean that it is morally permissible to engage in them. This way of reading him may help us to understand why he views enrichment as a pragmatic but not a moral concern.

It also, of course, may help to make sense of the justificatory approach he does take with respect to responsibility. Demands of fairness, after all, seem to be the domain of deontology.⁹ Further, Wallace seems primarily concerned with the mere permissibility of responsibility practices understood in terms of the standards that we already all clearly accept as limitations on appropriately

⁸ Wallace (1994, 102)

⁹ Whether this is really what many incompatibilists really have in mind would seem to be a fair question, though it is not one I will seek to answer. The way in which the incompatibilist charge is formulated and understood does considerable work for Wallace. Whether this charge, which has been made in many forms by many theorists with diverse theoretical commitments, ought to be treated as unified in this way would seem to be a legitimate point upon which one might press Wallace.

holding one another responsible. “Claims about the conditions of responsibility must be treated as hypotheses in normative moral theory and be subjected, accordingly, to the standards of argument that prevail within such theory.”¹⁰ He goes on, “I will understand moral principles to be propositions that isolate the features of situations or actions that justify particular claims of moral right; they specify the reasons that given course of action is *obligatory or prohibited*.”¹¹ I take it that the sort of case he has in mind—at least the sort that would be of interest to his incompatibilist interlocutor—is that we think it impermissible to hold children responsible for their misdeeds (at least in the same way we do adults). Substantive moral judgments like this are the kind that will bear on the justification of our moral responsibility practices, and we can see how the incompatibilist worry that responsibility practices are *impermissible* is not necessarily defused if we understand these practices merely as a way to enrich our experiences.

Setting aside whether it is legitimate to delineate what amounts to a moral claim and what amount to a pragmatic claim in this way, I propose we reconsider the sense in which responsibility practices may “enrich” our lives. On my reading, the point at hand is a deep one about how we value the strains and triumphs of relationships. One way to meet the incompatibilist’s normative moral charge is not merely to argue that the principle of fairness *permits* us to hold people responsible such that the practice is permissible, but that there is something we value about being held responsible such that these practices are in some sense *obligatory*. In other words, there is a sense in which we owe it to each other to hold one another responsible.

The implication of such a picture is that we in some sense reasonably demand that others hold us responsible. I propose to make sense of this in terms of the general commitments we demand of other people that in turn give rise to moral responsibility practice. These practices may

¹⁰ Wallace (1994, 109)

¹¹ Wallace (1994, 109)

yet be rendered just—even in the face of determinism—if they are part of a framework of commitments we are prepared to endorse in ourselves and demand of others. Strawson’s remark about the enrichment of human life, then, can be understood in terms of a higher-order commitment to the interpersonal commitments that involve being held and holding others responsible.

If all this is right, the nature and value of these commitments warrant scrutiny. How can we understand the commitments of which responsibility practices are a part and why do we think these commitments are valuable? My aim in pursuing this line is to call into question the relationship between the reactive sentiments and the conception of responsibility we endorse from a higher-order standpoint. If the reactive attitudes do not have the right sort of relationship to the first-order commitments that we endorse from a higher-order standpoint, then we can see that an adequate account of responsibility should not rely upon them.

1.2 | *The Value of Responsibility*

Why is it the case that we judge being held responsible as much a privilege as it is a burden?¹² And why do we think it worthwhile to make ourselves vulnerable to others’ betrayals and the accompanying setbacks and disappointments, especially in light of our collective, well-documented and all-too-familiar propensity to let others down and be let down by others?

I submit that our proclivity for holding and being held responsible is part of a commitment to cultivating reciprocal relationships. In turn, our endorsement of responsibility practices cannot be separated from this commitment. These relationships are the context in which we can exercise our moral powers to make claims on others and to be moved by the claims that others make on us. What these relationships require, however, is that we regard ourselves and one another as capable of exercising these powers—in other words, that we hold ourselves and others responsible.

¹² The phrase is Angela Smith’s (2005, 269).

Responsibility practices, on this picture, both spring from and demonstrate our commitment to relationships of mutual regard. It is in the context of this commitment that we endorse these practices.

1.3 | *Relationships & Responsibility*

I begin with an example that I take to illustrate an individual's reasons for endorsing responsibility practices. As one might expect, sitcoms are replete with examples of people navigating responsibility practices. This one comes from an episode of *Modern Family*.

Modern Family | Phil has an important presentation for work. He solicits the assistance of his mother-in-law, Gloria, who is to act as a 'plant' in the audience and ask a question that sets Phil up for his grand finale. The day of, she decides to indulge in a salon visit and in so doing risks missing the presentation. She fails to show up on time and as a result the presentation is an unmitigated disaster. Though she let him down, Phil initially does not show any resentment toward Gloria. Gloria, in turn, becomes angry with him. Phil might not be yelling; but she is. "Why aren't you angry? Why don't you yell at me?" she implores. Later she says, "That's how you know they love you, when they're comfortable enough to yell at you."¹³

Why, we might initially wonder, does Gloria *want* Phil to yell at her? She provides the answer herself: Gloria considers his anger in this circumstance to be essential to their having the sort of relationship she craves. Her final comment tells us that she views a willingness and disposition to hold her to account as implicated in the standpoint of someone who values a certain type of relationship: Phil should become angry with her if he is committed to a loving relationship in the same way that she is. Conversely, the failure to hold her to her misdeed would indicate he does not even view her as capable of acting in the way that their relationship requires or that he does not care whether the relationship proceeds in that fashion.

In some cases, not to blame someone who has done wrong implies viewing her from Strawson's 'objective standpoint.' Instead of treating her as a free and rational agent, I regard her as I might "a child or a pet."¹⁴ This obviously does not quite apply to Phil and Gloria. But it does tell us

¹³ *Modern Family* (2012)

¹⁴ Korsgaard (1996, 196)

something about the value of responsibility. The failure to hold someone responsible forecloses the possibility of a relationship characterized by trust and reciprocity. If I do not view you as capable of being moved by the reasons I give you in virtue of my role or relation with respect to you, then you typically do not have the chance to cultivate this capacity.

We may note that this applies across interpersonal relationships, including what might be called ‘the moral relationship.’ If I do not view you as capable of acting toward me in accordance with the moral reasons given to you by my very personhood, the possibility of a relationship based on these reasons goes out the window. Likewise, if I do not treat you as capable of being moved by the principles that govern, say, a loving familial bond or ‘perfect friendship,’ I foreclose the possibility that we are or can be in the desired sort of union.

Crucially, by failing to hold you responsible, I may also appear to waver in my own commitment to relating to each other in the right sort of way. If it matters to me that our relationship is conducted on terms of mutual respect, I will be disposed to treat you as though you are capable of acting on a similar commitment even when you have failed to do so. Insofar as reciprocal relationships require regarding one another as possessing moral and interpersonal capacities, not to hold one another responsible seems to indicate the absence of the right sort of commitment. Cultivating this sort of relationship requires regarding others as capable of acting on their own commitment to mutual regard, and regarding others in this way requires being disposed to hold them responsible.

1.4 | Commitment, Demands and Trust

Many theorists characterize reactive attitudes in terms of an essential relationship to the demands we make of other people.¹⁵ With this in mind, we can examine how demanding something

¹⁵ See Wallace (1994) and Darwall (2006, 74-9). A standard way of formulating this thought is that demands amount to a proneness to the reactive attitudes and that these attitudes (therefore) can amount to an expression of moral demands.

from someone involves a form of trust. Making a reasonable demand involves a belief that the target of one's address is capable of the sort of conduct being demanded ("ought implies can"). In moral and interpersonal cases, we can think of the capacity at issue in motivational terms. When I demand that you comport yourself in a certain way toward me, I simultaneously imply that you are liable to be moved by interpersonal reasons or by the principles that govern our relationship. In these moral and interpersonal cases, this amounts to a form of trust: I trust that you share my commitment to the value of a relationship in which we each have standing to make demands upon one another.

Ideally, we demand nothing of those whom we do not in some sense trust. I will not ask you to do something I believe you cannot, unless my intent is to shame or manipulate you. While we think it is unfair when, say, a boss asks us to do something impossible, we do not think it is unfair when people in our lives demand that we be moved by moral considerations, even if it turns out we cannot muster the will to act accordingly. Instead, we may be dismayed when we find out that little has been expected of us. "Don't you trust me?" is typically leveled as an accusation because we think we ought to be afforded the opportunity to prove our worth as moral agents (or, as the case may be, friends, family, lovers, etc.).

We are willing to be held responsible for our misdeeds because we generally endorse what it indicates about someone else that they are willing to hold us responsible. We can think of the trust implicit in demands as an indication of a commitment to a relationship conducted on reciprocal terms. Only if I am so committed will I be disposed to take the risk, so to speak, implicit in interpersonal expectations and demands. Only if I think that you and I, together, are the sort of agents who can conduct the right sort of relationship, does this sort of risk-taking make sense. To be subject to interpersonal demands is to be treated as one capable of being responsive to such demands in the context of a relationship. It also indicates that the other party is committed to

relationships of reciprocity insofar as she has taken the risk of trusting via expectation or demand.¹⁶

If we return to *Modern Family*, we can see all of this in action. In Gloria's mind, if Phil is not willing to hold her accountable for her flub, it means he is not invested in the sort of loving relationship she wants. This allows us to make sense of what she means when she says she wants him to be "comfortable" yelling at her. The comfort in question is the willingness to make the trusting leap implicit in a relationship-based demand. Not only does Gloria want Phil to regard her as capable of being moved by the ideal of a loving relationship, but she wants this sort of relationship to matter to him in a way that will dispose him to make demands of her.

Recall the question we set out to answer. What do this case and others like it tell us about the reasons we have for endorsing responsibility practices? Being held responsible means being treated as capable of a relationship based on reciprocity. Being treated as such is required for this sort of relationship. Though it is often unpleasant to be held responsible for our wrongdoing, we hope, in a general sense, that we will be. Not to be held responsible would foreclose the possibility of carrying on relationships that involve the actualization of our interpersonal capacities, a possibility that most of us, including Gloria, are eager to avoid. Further, others' willingness to hold us responsible says something not only about how they view us but about their own commitments. On this picture, that we value being held responsible is an essential part of our fundamental interest in realizing our moral and interpersonal capacities in the context of a reciprocal relationship.

1.5 | *Relational Morality and the "Problem of the Stranger"*

I have argued that responsibility practices are justified in virtue of their connection with a certain way of regarding one another and that the pattern of regard in question is constitutively

¹⁶ Korsgaard puts the point beautifully: "To hold someone responsible is to regard her as a *person*—that is to say, as a free and equal person, capable of acting both rationally and morally...Abandoning the state of nature and so relinquishing force and guile, you are ready to share, to trust, and generally speaking to risk your happiness or success on the hope that she will turn out to be human" (1996, 189-90).

related to a commitment to the value of conducting relationships on reciprocal terms. Generally speaking, we *want* to be held responsible because of what it tells us about the commitments of those around us and how those people view us in light of those commitments. But by cashing out the relevant commitments in explicitly relational terms, I open myself to an objection that has been raised against other relational accounts of responsibility and blame. Such accounts are not, of course, all identical, but share some relevant features that warrant some pushback. Generally speaking, the thought is that wrongdoing negatively impacts the relationships in which we stand to one another and that this fact makes it either permissible or obligatory to attempt to *mark* or *repair* this damage through responsibility practices such as blame or forgiveness.¹⁷

Now, it is not clear to me that what I've said to this point has exactly this shape. Nonetheless, I may be open to a form of the objection that George Sher makes to relational accounts. What, Sher asks, can these accounts tell us about people with whom we apparently have no relationship? According to relational theorists, "the damaged relationship is said to hold the key to understanding why the relevant response is warranted." But such an account must be silent in cases where we are wronged by a stranger—"by definition, someone with whom one has no relationship."¹⁸

The analogous form of this objection applied to my account is simply to wonder whether we really feel the same way about being held responsible by strangers as we do about being held responsible by those with whom we have particular personal relationships. The question of why we want to be held responsible in light of the pain and inconvenience of it becomes more pointed in such cases. Would we really not just prefer that others were systematically indifferent to our misdeeds? It's one thing to claim that *Modern Family's* Gloria hopes that her son-in-law cares about

¹⁷ See Radzik (2016, 107).

¹⁸ Sher (2013, 55)

their relationship such that he holds her to account. But when it comes to our “relationships” with strangers, perhaps we would simply prefer always to, well, *get away with things*. And if we do, then the justificatory structure I have proposed for responsibility practices seems to come apart when it comes to being moral responsible with respect to strangers.

I hasten to note that one reason we might be inclined to think that we do not have the same view of the value of being held responsible by strangers is that there are fewer things for which we think it is appropriate for a stranger to hold us responsible. Special kinds of personal relationships open up the possibility of new kinds of wrongs. That is to say, I can wrong my mother or my best friend in ways that it would be impossible to wrong a stranger. Were an unfamiliar passerby to impeach me, “You haven’t called me in a month!” I could concede the point but ask, fairly, “And why should I have?” I could not, of course, respond in the same way to my mother. By this I simply aim to point out that there are far fewer things for which we would want to be held responsible by strangers because there are far fewer things that we owe to strangers. So one possible response to the objection at hand is that it mistakes a difference in degree for a difference in kind: it’s not that we do not value the responsibility practices of strangers because there exists no possible moral relationship in which we might stand with them. Instead, we simply think there are not so many things that it would be good to held accountable for by such people.

This reply, however, is hardly decisive. Another possible response is that, for one reason or another, we are more keenly aware of the painful aspects of being held responsible in cases with strangers. This speaks directly to the issue at hand. The thought is that perhaps the anonymity of stranger cases means when we consider whether we want to be held responsible in them, we focus more on the undesirable dimensions of such a practice than the ones that make us endorse it in cases where there is some special relationship at stake. In this way, our judgment is “clouded,” in a sense; we cannot tell what we “really” think is best.

Now, talk of what we “really” judge best may seem dubious, especially when our supposedly authentic judgments are in open conflict with what we seem often to prefer on a case-by-case basis. But it is important to note that the standpoint from which we are making the sort of judgment I need is a general sort of one. The larger question at hand is about whether and how the practice of holding each other responsible for apparent misdeeds is justified. My answer involves the fact that we generally want people to have the sort of commitments that involve viewing us in a way that gives rise to responsibility practices. That we wish we could get away with things on a case-by-case basis does not seem to undermine this picture. As an analogy, consider the fact we think it is a good thing that we usually have to work for our success even though in any given case we might prefer to have certain things handed to us.

Another way to bring home my point about what we “really” want is to consider a figure who features disproportionately often in the responsibility literature: the psychopath.¹⁹ The first thing the psychopath offers us is a character who views others’ tendency to hold him responsible as a mere inconvenience. Of course, viewing others’ practices in this way is not sufficient for psychopathy. This mindset does, however, fall necessarily out of the psychopath’s moral illiteracy and inability to see the interests and claims of others as valuable and legitimate. I do not mean to make an argument of the form “You do not want to be a psychopath; therefore, you must want to be held to account for your misdeeds.” I do, though, want to draw attention to what we think about the life of the psychopath. It is, as Strawson says, “impoverished.” And this is so in no small part because he is missing out on those things that make us think there is value in being held responsible: shared commitments with others and the potential for relationships conducted on the terms of these commitments. Insofar as we think the psychopath is missing out on something we think is valuable, it seems that we “really” want to be the sort of people who can acknowledge the value of being held

¹⁹ For an iconic treatment of this topic, see Watson (2004a).

to account, lest we miss out in this way. But all that is simply another way of saying that we “really” want to be held responsible, even if there are times when we’d really rather get away with things.

The psychopath is instructive not only if we imagine the world through his eyes, but also when we imagine ourselves as the objects of his sterile gaze. This character cuts a disturbing figure not merely because of his potential to commit heinous acts and tendency unflinchingly to defy moral norms, but simply in virtue of how he views us. Even when the psychopath does not *do* anything *wrong, per se*, we are deeply put off by the possibility that he does not see us in the way required for the undertaking of a certain type of relationship. Even when he is not in a *position* to harm us directly, we feel this way. The sterile gaze of the psychopath is itself upsetting. For any apparent act of holding responsible in which he would be but a pantomime of the real thing.

By considering the psychopath, we can see that we need not be any particularly intimate sort of relationship with a person in order to harbor the hope that she regards us in a certain way. Nor is this hope contingent on the fact that if she *doesn't* regard us in this way, she might be disposed to harm us in some way. Instead, we hope even the stranger passing on the street views us not as she views the fire hydrant she must step around, but as the type of agent we see ourselves as. Our preference is for a world in which other people see something of themselves in in us: they recognize us as rich centers of values and beliefs, but also crucially as capable of reciprocating this way of viewing those whom we encounter. We want others to conceive of us as prepared to meet on shared ground to the extent that we are mutually committed to a certain way of regarding one another. But this is simply for them to conceive of us as possible participants in a shared sort of project, even if it is simply the project of relating in a way that is valuable to both of us. And such a world is one in which those around us—strangers and intimates alike—are poised to hold us to account for our misdeeds.

2.1 | *Endorsing Commitments vs. Endorsing Emotions*

The plausibility of the reactive account relies on there being a tight connection between proneness to the reactive attitudes and a commitment to cultivating reciprocal relationships. One way to understand this connection is that so long as someone has the relevant commitment, she will be prone to the relevant attitudes. This is the view that an affective dimension is part or constitutive of certain value judgments.²⁰ It is notable, however, that we can pull apart sentiments and commitments, at least conceptually. This is particularly apparent when we view things from the higher-order standpoint from which I recommend we interrogate the value of responsibility and the reactive attitudes.

Consider a powerful analogy presented by McKenna.²¹ In defense of the value of the harms we inflict upon someone when we blame her for wrongdoing, McKenna invokes the harms of grief. He suggests that the emotional pain we experience in grieving the death of a loved one is valuable because it is an expression of the value of what we have lost. True enough. But the conceptual distinction between what we value and our affective response to damage to it becomes clear if we examine our higher-order commitments here. If I ask you, “Would you like to be type of person who didn’t cry at your loving mother’s funeral?” you would surely respond in the negative. And if I asked “Why?”, you would likely give an account of how failing to have a certain emotional response would mean you never really cared for your mother in the first place. You would rather be the sort of person to whom a loving relationship with a parent mattered deeply.

But when we run these questions the other way, our higher-order commitments become apparent. If I ask, “Would you like to be the sort of person to whom your relationship with your

²⁰ See Wallace (2011) and Scheffler (2011).

²¹ McKenna (2017)

loving mother meant nothing?” it would be very odd for you to say, “No! Because then I wouldn’t cry at her funeral.” You do not want to be the sort of person who cries at funerals, *per se*; you want to be the sort of person who is invested in tender familial ties. Of course, it is natural to think such an investment involves being vulnerable to certain emotional states. But our reflective endorsement appears to be of the investment itself, rather than its emotional trappings.

With respect to moral emotions, this conceptual distinction is significant. A reactive account of responsibility would be highly plausible if the attitudes in question could not be pried apart from our commitment to relationships in which we hold one another responsible. Were this the case, the explanation for holding one another responsible in this way would be that these emotions arise in virtue of our commitment to moral and interpersonal expectations. But as things stand, it seems an account of responsibility that we are positioned to endorse must first speak to the right sort of interpersonal commitments and need not say anything about our emotions.

2.2 | *Why the Reactive Account Needs Reflective Endorsement*

Even if no such gap between the commitments that we endorse and the attitudes to which they purportedly give rise can be established, we should note something about the approach reactive account theorists take to justifying the more unpleasant aspects of these attitudes and their expression. Specifically, these arguments rely on certain commitments on the part of the reader. When McKenna argues that there is value in grief, he takes for granted that we want to be the sort of people who care about loving relationships. Likewise, arguments aimed at justifying the unpleasantness associated with guilt and blame assume that we think it is a good thing that we are people who care about their relationships with others such that the negative appraisal of those around us gives rise to this unpleasantness.

The question of why we think it worthwhile to subject ourselves to the burdens of responsibility is sensible to us at least in part because we are familiar with the unpleasantness of

being blamed for wrongdoing. This unpleasantness has diversely been termed the force, harm or ‘sting’ of blame.²² The force in question, I take it, is an all-too-familiar phenomenon. We need only think back to a time when we have been called out for breaking a promise, telling a lie or otherwise displaying disregard for someone we care about. Such reflection tends to elicit a painful pang, even if the offense is distant and any animosity has been patched over.

Notably, our distaste for being blamed does not seem merely to be a distaste for the consequences thereof. It is true that I may fear some form of sanction or be perturbed by the loss of some instrumentally valuable social capital. But being held responsible for wrongdoing is unsettling in a way that runs deeper than all this. Blame weighs upon us in its own right, not merely on the terms of the loss of something instrumentally valuable. We experience moral criticism in terms of the loss of something valuable in itself.²³

We can account for this by observing how deeply it matters to us that we eschew actions that cannot be justified to those around us. Being blamed marks a failure in this respect. Even in cases where we feel we have not been blamed appropriately—perhaps we think we have done nothing wrong or cannot quite figure out what line we crossed—we sometimes experience the same sort of force of blame. For when this happens, we are nonetheless faced with the prospect of irreconcilable differences in values which present as obstacles to treating one another in a way that is mutually intelligible and justified.

The force of blame is one reason philosophers have thought responsibility practices demand justification. The prevailing view of this justification is stated succinctly by McKenna. He tells us, “one who had no regard for the well-being of her interpersonal relations with others would not take damage to them as a detriment to her interests...Here the goodness is located in the blameworthy

²² On this topic, my thought has largely been influenced by Hieronymi (2004) and Scanlon (1998, chapter 6). See also McKenna (2012, chapter 7), Wallace (1994, 52-62) and Watson (2004b) for further discussions.

²³ Scanlon (1998, 271-2)

agent’s commitment to membership within the moral community.”²⁴ In the same vein, McKenna observes that in blaming a wrongdoer, “those who blame display their commitment to the moral considerations undermined by the blameworthy person’s wrongdoing.”²⁵

All of the foregoing fits in neatly with everything I have said to this point. Being blamed indicates that you have failed to reciprocate the trust placed in you by another person, thereby undermining the possibility of realizing the sort of relationship you value so highly. Because mutually respectful relationships with other human beings, and the ability to conduct and maintain these relationships, are among the things that we characteristically hold most dear, blame marks a setback to one of our most profound interests.

But note: whether this argument succeeds as a defense of the harms of blame depends on whether we agree that it is a good thing that people care about reciprocal relationships in this way. By no means do I want to suggest that we generally do not or should not agree. What I do want to point out is that arguments for a theory that claims to put a commitment to interpersonal standards on all fours with proneness to the reactive attitudes nonetheless appears to require first and foremost that we have reason to endorse this sort of commitment if we are to think of these attitudes as justified. The claim is that the reactive attitudes manifest a commitment to the principles that govern the moral community. But as in the example of grief, clearly what is more important to us—and for a defense of responsibility—is that people are committed to reciprocity.

§3 | CORRUPTIONS OF RESENTMENT

3.1 | *Interpreting Reactive Attitudes*

So far, I have suggested that we have a higher-order commitment to caring about reciprocal

²⁴ McKenna (2012, 167)

²⁵ McKenna (2012, 168)

relationships and that this endorsement is what justifies responsibility practices. I raised doubts about whether the commitment we endorse has any necessary connection to the reactive attitudes such that if we are so committed, we are prone to these attitudes. This is important because it suggests that the endorsement of certain commitments does not necessarily equate to the endorsement of the attitudes in question.

However, it may remain open to the reactive account theorist to suggest that inherent to these attitudes is the right sort of commitment. Strawson's well-known line on this point is that when we experience a reactive attitude toward someone, we "continue to view him as a member of the moral community; only as one who has offended against its demands."²⁶ Wallace supplies something a bit more robust in saying the reactive attitudes are distinctively characterized by their relationship to expectations and demands. He heads off an interpretation of the reactive account as crudely noncognitivist and takes pains to emphasize that the moral emotions are not "arbitrary feelings" but rather connected essentially to moral obligations.²⁷ And Stephen Darwall explicitly argues for the claim that the reactive attitudes are a form of moral address "which presupposes another's competence and standing to be thus addressed."²⁸ The general idea is that the reactive attitudes implicate a commitment to treating others as morally capable agents, able to be moved by the demands of respectful relationships.

Recall, though, that an essential part of the reason we think it valuable to have demands made of us is that we suppose the willingness to make demands is itself rooted in a commitment to reciprocal relationships. But even if our attitudinal responses to perceived wrongdoing generally reflect regard for another's interpersonal capacities, there is reason to doubt whether they necessarily do so as part of the right sort of commitment. To sow the seeds of this doubt, we need look no

²⁶ Strawson (1982, 34)

²⁷ Wallace (1994, 77-8)

²⁸ Darwall (2006, 70)

further than the uglier features of our practices. If essential to the reactive attitudes is an appreciation for the moral capacities of others and an interest in relating to them on reciprocal terms, why is it the case that these attitudes often come packaged with toxic animus and vindictiveness?

In response to this question, the reactive account theorist can say that buried deep in even our most manifestly rancorous reactive attitudes is some kernel of the recognition required in a genuine practice of responsibility.²⁹ But this point would be rendered moot if, within the standpoint from which we experience the reactive attitudes, we did not derive the requisite presuppositions about others' interpersonal capacities from our own commitment to reciprocal relations. To this point, I will argue that the reactive attitudes do not necessarily carry the right sort of implication with respect to the commitments of those prone to them to serve as a satisfactory basis for an account of responsibility.

3.2 | *Corruptions of Resentment*

In an intriguing passage, Darwall argues that the pure form 'self-conceit' required to untether the reactive attitudes from an assumption of others' moral capacities in their object is vanishingly rare and arguably pathological.³⁰ To support this claim, Darwall invokes Stalin, a man we might think as liable as any to the delusions of self-conceit, but biographical details of whom reveal as "someone who was motivated by richly elaborated reactive attitudes, which, to be sure, he marshaled for his own ends." Of Stalin, Darwall writes,

His emotional life was replete with episodes, frequently staged, in which a justified authority over others seemed manifest to him, justified in ways that, as it seemed to him, others should be able to appreciate. And even his cruelest murders were accompanied, indeed fueled, by self-justifying emotions and narratives. It seems no exaggeration to say, in fact, that Stalin's distinctive form of evil essentially employed a cynical and distorted form of moral self-justification that he manipulated for

²⁹ Darwall defends this point (2009, 77).

³⁰ Darwall explains Kant's notion of self-conceit as "a fantasy about second-personal status" in which a person considers his own will to be uniquely normative for others. "A thirsty person with self-conceit will take it that others have reason to relieve his thirst because this is what he wills or wants (though he would have no such reasons if roles were reversed)" (2006, 135).

his own purposes. Thus, Radzinsky [Stalin's biographer] tells us, a standard ploy of Stalin's was to catch someone he wished to "eliminate" in a lie, after which Stalin "felt entitled to feel a moral hatred for the liar and traitor."³¹

Darwall's point is that even despots do not view their own expectations for others as absolutely binding, *a la* divine commands. Instead, they justify their vicious and self-serving attitudes and conduct as part of an appropriate response to others' misdeeds. By couching hatred and domination in righteous resentment, Darwall says, Stalin's perverted reactive attitudes nonetheless involve recognizing the ability of his victims to appreciate and respond to interpersonal reasons.

In my view, Darwall's analysis misses something crucial. Stalin's reactive attitudes are, at bottom, measures of self-justification. They begin and end with his contempt for others and are aimed primarily at confirming what he already believes about himself in relation to other people. The curious thing about Stalin's attitudes is that they are largely about Stalin. Only incidentally do they involve the conduct and capacities of their objects. They are responsive to others' perceived misdeeds or ill will only insofar as reacting to these things is useful to reinforce the way he is already disposed to treat them. His reactive attitudes, then, are not part of a genuine commitment to any sort of reciprocal relationship; they are first and foremost a way to justify attitudes he already has towards those around him.

The perversion in Stalin's case is profound. But lesser forms of it would seem to be all-too-familiar. We can imagine an emotionally abusive boyfriend who summons genuine resentment at his girlfriend's minor, perceived slights in order to justify his pre-existing scorn. If he achieves his aim of inciting guilt in his partner, he may even succeed in making her complicit in his contempt, causing her to feel as though she has failed to hold up her end of an arrangement in which they relate on mutually respectful terms. In such a case, as with Stalin, the reactive attitudes may indeed be predicated on the recognition of others' moral capacities. The man in question flexes his normative

³¹ Darwall (2006, 139). Darwall refers to Radzinsky (1997).

muscles, so to speak, in part because he believes that his partner will give his resentment uptake. But he does so in bad faith. In these cases, instead of viewing these capacities and the person characterized by them in terms of the potential for a reciprocal relationship, they are exploited—consciously or otherwise—to justify certain affective and practical stances towards others or to paper over one’s own deficiencies.

Here, it may be helpful to consider the failure of consequentialist theories of responsibility practices as a point of comparison. From Strawson onward, advocates of the reactive account have emphasized its advantages over deterrence-based models that paint responsibility practices as a way to regulate others’ conduct via their fear of the unpleasant consequences generally associated with being held responsible for wrongdoing. This approach is at first blush a rather natural way of accounting for the relevant attitudes and practices which says that while we might harm someone when we blame him, that is part of the point. Blame and its associated harms deter people from wrongdoing because they do not want their harms returned, much less returned in kind. The practices are then justified in terms of their ability to produce desirable outcomes by deterring bad acts.

Such a picture has a number of problems.³² One complaint concerns the unsavory implication that we are justified in sanctioning someone who has done nothing wrong insofar as doing so is a good way to prevent future bad actions. This objection strikes me as symptomatic of a deeper concern. The deterrence model is incompatible with the justification we have for endorsing a framework for responsibility insofar as it implicates a failure to regard one another in terms of the capacity to be responsive to moral and interpersonal reasons.

If our endorsement of responsibility is due to the essential connection it has to a

³² Wallace succinctly discusses three such problems (1994, 54-9). See also Ebels-Duggan (2013, 146-7n) for a concise treatment that helped me clarify my thoughts on this matter.

commitment to mutual regard, regulating someone's conduct in what Wallace calls an "economy of threats" cannot amount to holding her responsible.³³ If I am not willing to give you the right sort of reasons to comply with the terms of our relationship, opting instead to treat you as a force to be managed, it appears I am not really committed to the right sort of relationship myself. Instead of emphasizing others' capacities to understand and act from interpersonal reasons, responsibility practices in an economy of threats exploit the economic rationality of mature adults in order to regulate their conduct. By substituting compliance and enforcement for commitment and endorsement, the deterrence model seems not to be part of a commitment to reciprocal relations at all insofar as it gives others the wrong sort of reasons.

It is tempting to think that the reactive attitudes' implicit presupposition of the ability of others to respond to our normative powers means the reactive account can avoid this worry. We have just seen, however, that while resentment and its ilk may indeed carry this sort of supposition with respect to others' commitments and capacities, these attitudes do not always have the same implications for one's own values. To the contrary, it is not hard to imagine cases of resentment that serve basically to justify attitudes antithetical to reciprocal commitments, such as disdain and self-conceit.

In denying the reactive account, I do not mean to suggest that we never respond to wrongdoing or hold others to account when we experience and express these attitudes. As with the deterrence model, this would go too far: clearly, we often have the deterrent effects of blame in mind when we hold others responsible. But in both cases, the accounts on offer fail because the pictures they paint allow space for holding responsible from a standpoint other than the one we endorse as the justified basis for this practice. The standpoint we adopt when we aim to deter is regulatory rather than collaborative. Likewise, the standpoint from which we experience the reactive

³³ Wallace (1994, 54). Wallace borrows the phrase from H.L.A. Hart.

attitudes is not one that necessarily includes a commitment to relating on reciprocal terms, but instead one that at times is based on an interest in justifying one's own failure to be so committed.

§4 | RESPONSIBILITY & MODALITY

4.1 | *Vulcans, Cheerleaders & Valuing*

I have suggested that the reactive attitudes are not essential to moral responsibility because they do not arise necessarily from the commitments that *are* necessarily implicated in responsibility. But at this point, the reactive account theorist can say that he need not take so hard a line on the nature of this relationship. McKenna, in fact, makes what he calls a “major concession” when he suggests that Strawson went too far in “rigidly designating” our responsibility practices. We cannot rule out the possibility of, say, a race of affectless Vulcans who succeed in holding one another responsible absent any liability to resentment and its ilk.³⁴

Derk Pereboom, for one, take this possibility seriously. He says,

it seems possible to imagine rational but emotionless beings who yet have a deep concern for right and wrong, and who believe that agents are morally responsible. Such beings would believe wrongdoers to be morally responsible without having any emotional attitudes, such as indignation or moral resentment, towards them.³⁵

The thought is that even if our actual responsibility practices turn on moral sentiment (or tend to), there is nothing necessary about them doing so. This would seem to be a problem for the reactive account if its aim is to provide necessary and sufficient conditions for moral responsibility.

Now, this sort of argument may strain credulity. Is it really possible for us genuinely to value a certain type of relationship without being prone to relevant emotions when someone fails to act in the way such a relationship requires? More generally, we might wonder whether it is possibly really to value anything at all without vulnerability to certain emotions related to the object of interest.

³⁴ McKenna (2012, 111-3)

³⁵ Pereboom (2001, xx-xxi)

Samuel Scheffler suggests as much, saying that valuing some item involves not only deliberative considerations—i.e., viewing the item in question as reason-giving—but also being emotionally vulnerable in some way with respect to that item. Part of valuing a hobby, for example, is to be upset when one is unable to engage in it. Part of valuing a friendship is to be distraught when it ends or to resent your friend she does something to damage the relationship. This emotional component, Scheffler says, allows us to differentiate between believing that something is valuable and actually valuing it.³⁶ I can appreciate the religious rites of an unfamiliar faith as valuable without finding them personally moving, for example.

This sort of interpretation is plausible and would suggest that, assuming we value relating to one another on moral terms, then we are necessarily emotionally liable to resentment and the like as a result. But while there seems to be something to this thought, the possibility that this affective dimension of valuing is contingent appears to remain a live one in light of the argument from reflective endorsement that I have already made. At least in the case of moral responsibility, that the reactive attitudes might be merely epiphenomenal seems like a possibility we should take seriously when we consider what, exactly, we demand of one another on moral terms. These demands are primarily for practical commitments. That is, they deal with the fact that we think people should view other people and certain types of relationships as reason-giving. Indeed, it seems a distinctive feature of moral reasons is that they have overriding deliberative weight. They are supposed to move us even in the absence of the influence of affect and desire and often in spite of the countervailing influence of these factors. So in terms of what we most basically demand of others in terms of a commitment to morality, the deliberative dimension of valuing is more basic than the affective one.

This is roughly the same point as the foregoing one about reflective endorsement. We hope that a commitment to morality will manifest in a person's conduct, not merely in her emotional

³⁶ Scheffler (2011, 30-1).

states. We can say something similar about how we react to wrongdoing. It is true that someone who values reciprocity and respectful relationships may as a result experience indignation when she is party to moral malfeasance. But it is not clear that such a reaction is really what we demand of her when we demand that she cares about moral conduct, and therefore not clear that her reactive attitudes are justified on their own terms. They are justified insofar as they are implicated in her moral commitments, which in turn are the commitments judge favorably and which therefore carry the justificatory weight with respect to responsibility practices.

McKenna, in any case, does not deny that that connection between emotional vulnerability and valuing is a contingent one. His response to the objection involves granting the claim that the connection in question is contingent while denying that it is much of a problem for reactive account theorists. To argue for this point, he presents an analogy to language. The suggestion that we could value without affect—specifically, practice responsibility without the reactive attitudes—is on par with the thought that “rather than use our voice boxes to make utterances, we could contort our bodies into the shapes of the different letters of some alphabet, spelling out each of our utterances one letter at a time like gyrating cheerleaders on amphetamines.”³⁷ The possibility is a live one. Yet when it comes time to concoct a theory of meaning, McKenna argues, we do no wrong by building our theory around the contingent manifestations of it with which we are most familiar (namely, human speech and written language).

This is somewhat convincing as it stands, although I think it is perfectly open to someone to resist by pointing out that while a theory of language must address these contingent matters, a theory of meaning should get at something a bit deeper. The fact that we think maybe we could convey meaning via spastic gyrations in much the same way we can through text and speech means we think there is some more significant common thread which can be run through cases of meaning and

³⁷ McKenna (2012, 112)

communication.

Beyond this, however, it seems the cases of language and responsibility are disanalogous. The concern in a theory of language, I take it, is largely descriptive or definitional. The inquiry is motivated by a theoretical interest in providing the necessary and sufficient conditions for certain concepts and phenomena. But the concern in a theory of responsible is different insofar as it is normatively loaded. We can trace this concern through the dialogue between compatibilists and incompatibilists. The query of the incompatibilist is not “what does it mean to be responsible, anyway?” Instead, she wants to know how we are *justified* in holding one another responsible. The concern is that if we extrapolate from the conditions that make holding someone responsible unjustified under normal circumstances (e.g., she is a child, he couldn’t have done otherwise, etc.), we begin to feel as though it is unjustified in any and all circumstances if determinism is true.³⁸

One way we can justify these practices, as I have already suggested, is if they are part of a set of commitments that we have reason to endorse. But this, I submit, gives us new reason to demand a necessary connection between certain commitments and the attitudes and practices that amount to a theory of holding responsible. Importantly, this reason is not grounded in a merely theoretical interest in providing an account with the definitional virtue of accurately delivering necessary and sufficient conditions for the concept in question. Instead, we have a moral interest in a theory that itself achieves the aim of attributing to us the commitments that make responsibility practices valuable. A theory according to which we are subject to certain affective states which are merely contingently connected to the set of values that we think make responsibility practices worthwhile is

³⁸ Nagel, perspicuously: “The problem of free will, like the problem of skepticism, does not arise because of a philosophically imposed demand for external justification of the entire system of ordinary judgments and attitudes. It arises because there is continuity between familiar “internal” criticism of the reactive attitudes on the basis of specific facts, and philosophical criticisms on the basis of supposed general facts. When we first discover the possibility that all human actions may be determined by heredity and environment, it threatens to defuse our reactive attitudes as effectively as does the information that a particular action was caused by the effects of a drug” (1986, 125).

inadequate. This modal imprecision amounts to a failure to emphasize what renders responsibility practices justified in the first place. It would be better if what made us responsible, understood in terms of being held responsible, was connected necessarily to what is good about being held responsible.

4.2 | *Theories of Better and Better Theories*³⁹

There may appear to be something suspicious this last argument. I have indicated that, because we think it would be better if an account of responsibility that necessarily implicated being held responsible as a result of others' commitment to respectful relationships was true, such an account is true. And as a result, I suggested that an account that centers reactions that do not arise necessarily from the right kind of commitments on the part of those who hold us responsible cannot be right. So the argument has an unusual form. It says that because it would be better if a certain account of responsibility was true, we should think it is preferable to the going theory.

To lend the argument an air of plausibility, I will begin by appealing baldly to authority and point out that ones like it have been made before. In "Personal Rights and Public Space," Thomas Nagel considers the possibility that we have inviolable rights because a theory in which we have such rights is preferable to a theory in which these rights are conditional in some sense. The thrust of the argument is that there is something inherently valuable about it being impermissible to be treated in certain ways. According to Nagel, "we can distinguish the desirability of not being tortured from the desirability of its being impermissible to torture us...To be tortured would be terrible; but to be tortured and also to be someone it was not wrong to torture would be even worse."⁴⁰

Ryan Preston-Roedder has explored this form of argument further.⁴¹ He suggests that we

³⁹ The title is inspired by a recent presentation by Geoff Sayre-McCord (2017), which addresses sort of argument I discuss below.

⁴⁰ Nagel (1995, 93)

⁴¹ Preston-Roedder (2014)

have reason to prefer theories in the moral realm that *themselves* do a better job with respect to some substantive moral aim when compared to competing theories. By way of example, Preston-Roedder says that having autonomy, understood as the capacity to conceive, fix and pursue one's own ends rather than have obligations and ends externally imposed, is such a moral aim. As such, theories that afford us this autonomy rather than simply supplying moral aims for us are preferable to, for example, act utilitarianism, which prescribes heteronymous ends (or *a* heteronymous end, as the case may be). Though Preston-Roedder does not cite him as an example, I think we might find a hint of this sort of argument in Rawls's contractualist attack on utilitarianism, in which he claims that parties in the original position would not opt for a society governed by principles that even allowed for the possibility they would be exploited in service of some agent-neutral conception of the good.⁴²

In this spirit, I submit that a theory of responsibility should bear a necessary connection to the substantive aim that practices of responsibility can achieve. This aim, and the reason we endorse and engage in these practices, has to do with the practical regard we have towards other rational beings. We suppose being and holding responsible is valuable because it arises from people's commitment to establishing and maintaining moral relationships. And we suppose the greatest excellence attained by these practices is in their role in the realization of this relational ideal. As such, we should prefer a theory of moral responsibility understood in terms of ways of holding responsible that necessarily implicate the right sort of commitments on the part of those holding others to account.

We hold one another responsible because we are committed to conducting mutually respectful relationships and because this commitment means we must think other people are committed to doing the same. Importantly, part of why we think it is a good thing to be held responsible even though it is often quite painful to be blamed for wrongdoing is that, in doing so, others indicate that they are committed to relationships conducted on moral terms. To hold

⁴² Rawls (1999).

someone responsible is justified, we think, because it springs from a valuable commitment to morality. In this sense, we are willing to be held responsible because being held responsible affords us a certain moral status: that of a capable collaborator in the moral community.

Compare the moral status of being responsible to the status of inviolability discussed by Nagel. In both cases, having the status in question opens us up to some form of hardship that would not be possible were we not to have this status. In the case of responsibility, we can be held to account for our misdeeds. We are justly subject to moral criticism, both from others and from ourselves, something that tends to be quite unpleasant. In the case of having inviolable rights, we of course open ourselves up to the possibility that we will be violated. If we did not think of ourselves as responsible for our conduct towards others or if we thought it was permissible for others to wantonly cheat and lie to us (for example), moral criticism and rights violations might not perturb us in the way that they do.⁴³ Yet it is clear on both fronts that we would not want to forego the moral statuses in question in light of the fact that they make us more liable to a certain sort of injury. With respect to inviolability, Nagel says that “this is just an instance of the fact that those who have more can lose more, and that it can be worth it.”⁴⁴

The same, it seems, might be said of responsibility. But we should be clear on what, exactly we think is “worth it.” Just as it means less to me for it to be wrong to murder me because doing so would decrease the average utility, so, too, does it mean less to me to be responsible for my actions in light of another person’s affective states. As such, the affective account gets us the wrong sort of responsibility. Responsibility is justified because we conceive of it in terms of the commitment to

⁴³ These things might of course still perturb us in some other way. Blame might involve the loss of some valuable social capital. Being cheated out of one’s money might make it hard to pay rent. But the point is that they would not amount to the same sort of injury we commonly associate with these things. It is because it matters to us that we respect others that being blamed carries the force it does, over and above the loss of social capital. And being duped costs more than the money you lose because it indicates others’ failure to treat you with due respect.

⁴⁴ Nagel (2008, 112)

morality. To value a certain type of relationships is to be disposed to hold another responsible. These are the terms on which we are prepared to endorse being and behind held responsible. But to dilute the claim by making the existence of phenomenon in question conditional on attitudes that are not necessarily part of this commitment is to nullify the value of the phenomenon in the first place. Just as a right to bodily integrity is devalued if it is contingent on my ability to contribute to society, so is the privilege of being held responsible devalued when it is defined in terms of contingent attitudes. We should, therefore, insist that responsibility be understood only in terms of that which is essential to the sort of commitments that justify it as a practice, rather than in terms of contingently associated emotions.

CONCLUSION

With all of these pieces in place, I would like to make the shortcomings of the reactive account explicit. In its simplest form, this view is that the reactive attitudes have an essential role to play in what it means to be and to hold one another responsible. Strawson's stroke of brilliance was to show us that the concepts of being and holding responsible can be cashed out in practical rather than metaphysical terms. But even if the relevant practices are unavoidable, incompatibilists may object that they remain unjust. At this point, it is open to those sympathetic to a view of responsibility as a practice to say that not only *is* responsibility a part of human life, but it *should* be insofar as it is implicated in some of what matters most to us. What matters to us here is that our relations with other people are characterized by mutual regard, and that we are committed to the value of relating to one another in this way. In light of all this, a satisfactory account of responsibility is answerable first and foremost to the need to capture a way of regarding ourselves and others that implicates a commitment to reciprocal relationships.

With respect to the reactive account, this means the reactive attitudes must bear some deep

connection to a commitment to regarding one another in this way. I have explored two ways of understanding the reactive account theorist's claim that the reactive attitudes are constitutive of this commitment. In §2, I noted the conceptual distinction between what we endorse when we endorse a commitment to reciprocal relationships and the emotional trappings of this commitment. In §3, I argued that though the reactive attitudes may involve the presupposition of others' moral and interpersonal capacities, there are notable cases in which this presupposition is not part of the right sort of commitment. In these cases, we do not view others' moral capacities through the lens of their value in maintaining reciprocal relationships. Instead, the reactive attitudes are deployed as tools of manipulation and domination or perverted to justify one's own disdain for reciprocity.

I should be clear on one final point. What ultimately is at issue is not simply the fact that we sometimes respond to one another's wrongdoing in ways that are inappropriately disjoint from responsibility's requisite regard for others' moral capacities, though this is doubtless true. The concern is that the reactive account would have us adopt a conception of responsibility that includes attitudes and practices that do not arise from the standpoint we endorse when we endorse responsibility. We are responsible to one another because we are invested in regarding one another on reciprocal terms. A conception of responsibility, properly construed, must bear a necessary connection to a commitment to reciprocal relationships because this is the commitment from which our disposition to hold others responsible arises. In light of this, a conception of responsibility that admits attitudes and practices that do not arise from such a commitment or involve the regard for others in terms of their capacity to understand and act from interpersonal principles must fail. Because the right sort of commitment neither entails nor is entailed by the reactive attitudes, the reactive account does not get at what is fundamental to practices of responsibility.

REFERENCES

- Arpaly, Nomi. *Merit, Meaning, and Human Bondage: An Essay on Free Will* (Princeton: Princeton UP, 2006).
- Darwall, Stephen. *The Second-Person Standpoint* (Cambridge: Harvard UP, 2006).
- Ebels-Duggan, Kyla. "Dealing with the Past: Responsibility and Personal History," *Philosophical Studies* 164 (2013): 141-61.
- Gaus, Gerald. *The Order of Public Reason* (New York: Cambridge UP, 2011).
- Hieronymi, Pamela. "The Force and Fairness of Blame," *Philosophical Perspectives* 18 (2004): 115-48.
- Korsgaard, Christine M. *Creating the Kingdom of Ends* (Cambridge: Cambridge UP, 1996).
- McKenna, Michael. *Conversation and Responsibility* (New York: Oxford UP, 2012).
- McKenna, Michael. "Punishment and the Value of Deserved Suffering." Presentation at University of Wisconsin-Milwaukee Philosophy Spring Colloquium. Milwaukee, WI, April 7, 2017.
- Modern Family*. "Egg Drop." Season 3, Episode 12. Directed by Jason Winer. Written by Paul Corrigan and Brad Walsh. ABC (2012).
- Nagel, Thomas. *The View from Nowhere* (New York: Oxford UP, 1986).
- Nagel, Thomas. "Personal Rights and Public Space," *Philosophy & Public Affairs* 24 (1995): 83-107.
- Nagel, Thomas. "The Value of Inviolability," in *Morality and Self-Interest*, ed. Paul Bloomfield (New York: Oxford UP, 2008): 102-13.
- Pereboom, Derk. *Living without Free Will* (Cambridge: Cambridge UP, 2001).
- Preston-Roedder, Ryan. "A Better World," *Philosophical Studies* 168 (2014): 629-44.
- Rawls, John. *A Theory of Justice: Revised Edition* (Cambridge: Harvard UP, 1999).
- Radzinsky, Edvard. *Stalin*. (New York: Anchor Books, 1997).
- Radzik, Linda. "Relationships and Respect for Persons," in *Reasonable Responses: The Thought of Trudy Govier*, ed. Catherine E. Hundleby (Windsor Studies in Argumentation, 2016): 105-27.

- Russell, Paul. "Strawson's Way of Naturalizing Responsibility," *Ethics* 102 (1992): 287-302.
- Sayre-McCord, Geoff. "On a Theory of a Better Moral Theory and Better Theory of Morality."
 Presentation at University of Wisconsin-Milwaukee Philosophy Fall Colloquium. Milwaukee,
 WI, September 15, 2017.
- Sher, George. "Wrongdoing and Relationships: The Problem of the Stranger" in *Blame: Its Nature and Norms*, ed. D. Justin Coates and Neal A. Tognazzini (New York: Oxford UP, 2013): 49-65.
- Scanlon, T.M. *Moral Dimensions* (Cambridge: Harvard UP, 2008).
- Scanlon, T.M. *What We Owe to Each Other* (Cambridge: Harvard UP, 1998).
- Scheffler, Samuel. "Valuing" in *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*, ed. R. Jay Wallace, Rahul Kumar, and Samuel Freeman (New York: Oxford UP, 2011): 23-42.
- Smith, Angela. "Responsibility for Attitudes: Activity and Passivity in Mental Life," *Ethics* 115 (2005): 236-71.
- Strawson, P.F. "Freedom and Resentment" in *Free Will* (1st ed.), ed. Gary Watson (Oxford: Oxford UP, 1982): 59-80. (Original work published 1962).
- Wallace, R. Jay. *Responsibility and the Moral Sentiments* (Cambridge: Harvard UP, 1994).
- Wallace, R. Jay. "Dispassionate Opprobrium: On Blame and the Reactive Sentiments" in *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*, ed. R. Jay Wallace, Rahul Kumar, and Samuel Freeman (New York: Oxford UP, 2011): 348-72.
- Watson, Gary. "Responsibility and the Limits of Evil" in *Agency and Answerability: Selected Essays* (New York: Oxford UP, 2004a): 219-59.
- Watson, Gary. "Two Faces of Responsibility" in *Agency and Answerability: Selected Essays* (New York: Oxford UP, 2004b): 260-88. (Original work published 1996).