

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Philosophy Dissertations, Theses, & Student  
Research

Philosophy, Department of

---

Spring 4-23-2015

## Blame within Reason

Adam R. Thompson

University of Nebraska-Lincoln, adamrosst@gmail.com

Follow this and additional works at: <http://digitalcommons.unl.edu/philosophydiss>



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

---

Thompson, Adam R., "Blame within Reason" (2015). *Philosophy Dissertations, Theses, & Student Research*. 10.  
<http://digitalcommons.unl.edu/philosophydiss/10>

This Article is brought to you for free and open access by the Philosophy, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Philosophy Dissertations, Theses, & Student Research by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

BLAME WITHIN REASON

by

Adam R. Thompson

A DISSERTATION

Presented to the Faculty of  
The Graduate College at the University of Nebraska  
In Partial Fulfillment of Requirements  
For the Degree of Doctor of Philosophy

Major: Philosophy

Under the Supervision of Professor Mark van Roojen

Lincoln, Nebraska

May, 2015

BLAME WITHIN REASON

Adam Ross Thompson, Ph.D.

University of Nebraska, 2015

Adviser: Mark van Roojen

My dissertation develops a novel response to global skepticism about responsibility—the view that no one is fit to be held responsible for anything. Though P.F. Strawson offered a highly influential account of holding and being responsible, his argument is widely considered to fail as a response to global skepticism. The primary worry is that he only *describes* our practice of holding responsible but does not *justify* it. I propose an unorthodox Strawson-style account of holding and being responsible and employ that account to offer an argument against global skepticism which not only describes but also justifies our practice of holding individuals responsible.

A key innovation of my account is that it rejects the Humean idea that evaluative beliefs cannot motivate absent aid from independent desire. Rejecting that idea allows us to preserve the beauty of Strawson’s original reactive attitudes account, which cites specific emotions as the only attitudes involved in holding responsible. I go on to develop the moral psychology at work which allows me to support the idea that holding responsible involves more than emotional responses to an individual’s behavior. This also grounds support for the idea that the community of morally responsible agents is likely wider than traditionally thought.

Furthermore, I immunize my account and argument to well-entrenched concerns for Strawson-style views. Often I achieve this immunity by adopting views that are thought to be incompatible with a Strawson-style argument. These views are typically embraced by skeptics or other anti-Strawsonian theorists. For instance, contrary to tradition, I show that a Strawson-style argument can be founded on the idea that experiencing individuals as natural, mechanistic beings is compatible with the attitudes involved in holding an individual responsible. I go on to show that the theses I argue for together with others allow us to accurately describe and justify our practice of holding individuals responsible.

for Gena

We are accidents waiting to happen.

- rh

## ACKNOWLEDGEMENTS

There are so many who deserve to be thanked. First, I thank the Department of Philosophy at the University of Nebraska-Lincoln (UNL) for financial and academic support. The courses and discussions I've had with the professors, visiting speakers, graduate students, and undergraduates at UNL were invaluable. Of particular note is Steve Swartzler whose kindness and generosity met me at the door and continues today. Likewise, Al Casullo's steady guidance helped cultivate a mature appreciation for rigorous philosophical inquiry. Though he might think he failed given my chosen dissertation topic, Al has never stopped cheering for my success. Finally, thanks to Mark van Roojen. Mark never takes enough credit for his awesomeness so I try to tell everyone about it when I can. Usually they say, "I know, he should chair your committee." I love smiling as I say, "He is!" I look forward to his ongoing mentorship.

And, thanks to my dissertation committee: Aaron Bronfman, Ari Kohen, Robert Schopp, and David Shoemaker. It has been a pleasure to work with these folks. I cannot show enough gratitude to Dave who without even meeting me in person agreed to join the committee. In short, Dave went way beyond the call of duty. I can only hope that my revisions reflect the depth and precision of his expert commentary.

I must thank all of those who attended my talks throughout the years and were kind enough to offer constructive feedback on my mostly confused ideas. Of special note are Patrick Arnold, Gabe Brugier, Dave Chavez, Janice Dowell, Aaron Elliot, Luke Elwonger, Allison Fritz, Cullen Gatten, Shane George, Christopher Gibilisco, Reina Hayaki, Landon Hedrick, David Henderson, Ben Henke, Andrew Jaeger, Tim Loughlin, Jennifer McKitrick, Chris Richards, Chelsea Richardson, Dave Sobel, Andy Spaid, Kaz Wanatabi, and Preston Werner.

Finally, I owe a debt of gratitude to my mom, dad, sister, and brother all of whom make an appearance in this dissertation and have never stopped supporting my dream. Speaking of family, I cherish the conversations I've had with Cliff Hill over yard work, car work, Star Wars games, and music about the topics of this dissertation, philosophy generally, and pedagogy. And, of course, without my wife Gena I would be lost. (Okay, Simi, thank you too.)

# Blame within Reason

## CONTENTS

1:	Introduction	1
2:	Normativism and Resentment	13
3:	Recognition and Resentment	46
4:	Restrictions and Resentment	75
5:	Self-Revelation and Resentment	99
6:	Negligence and Resentment	126
7:	Optimism and Resentment	153
8:	Conclusion	184
	<i>References</i>	215

## LIST OF MULTIMEDIA OBJECTS

Table 8.1	190
-----------	-----

## I: INTRODUCTION

The picture that I develop across the seven essays of this work is multidimensional in aim. Individually, each chapter is a self-contained paper insofar as the arguments of each do not rely on the main theses argued for in the others. Together, the individual essays develop and defend my unorthodox Strawson-style account of moral responsibility. In that vein, my unorthodox picture also serves to underwrite the final chapter's novel response to global skepticism about responsibility—the view that no one is fit to be held responsible for anything. Though P.F. Strawson offered a highly influential account of holding and being responsible, his argument is widely considered to fail as a response to global skepticism. The primary worry is that he only *describes* our practice of responding to individuals as responsible but does not *justify* it. The Strawson-style account I sketch in these essays attempts to go some way to do both.

To see how the self-contained essays contribute to a strong response to global skepticism consider the following argument for skepticism<sup>1</sup>:

### The Skeptical Argument

- P1) Responding to an individual S in way W as an agent morally responsible for  $\phi$  is appropriate only if S is fit (or deserves) to be responded to in way W for  $\phi$ -ing.
  
- P2) No one is fit (or deserves) to be responded to as a morally responsible agent (i.e. No one is the fit (or deserves to be the) subject of moral-responsibility responses). (Anti-Fitness Thesis)

---

<sup>1</sup> I thank a discussion with Matt King in which he coined the term 'moral-responsibility response' as replacement for the term 'holding responsible'.

- C1) Therefore, responding to anyone as a morally responsible agent is always inappropriate.
- P3) If moral-responsibility responses are always inappropriate, then no one is morally responsible.
- C2) Therefore, no one is morally responsible.

My main target in the end is (P2), i.e., the Anti-Fitness Thesis. But, concentrating first on (P3) takes us deep into the Strawson-inspired theories of moral responsibility. Hence, it will offer a good starting point to elaborate the contribution that each essay makes to my Strawson-style account and argument.

One main tie that binds those strongly influenced by Strawson is their commitment to the following fitting-attitudes account of holding responsible:

- (N) An agent, S, is morally responsible (for  $\phi$ ) if and only if it would be fitting to hold S responsible (for  $\phi$ ).<sup>2</sup>

Some Strawsonians go to lengths to establish the thesis that (N) is the *only* analysis of being morally responsible that can do explanatory work in our theories of moral responsibility.<sup>3</sup> In Chapter 2: *Normativism and Resentment*, I argue that they have not made their case. I show that we can nevertheless capture key Strawsonian claims that make his picture so plausible and incisive.

In the background of Chapter 2 is my inclination to accept a buck-passing account of normative responsibility and worthiness. Roughly, on that account, *being normatively*

---

<sup>2</sup> Wallace (1994, 91). I do not replace 'holding responsible' with responsibility-responses terminology so that it is easily recognizable to those who accept it.

<sup>3</sup> See, e.g., Wallace (1994, esp. Chpt. 4) and McKenna (2012, esp. Chpt. 2).

*responsible, being praiseworthy, and being blameworthy* is a second-order property with multiple first-order realizer properties. Briefly, then, the first-order realizer properties that make S morally blameworthy for breaking a promise can provide U with reason to blame S for breaking that promise.<sup>4</sup> Though that type of picture is in the back-ground, I do not lean on it here. The question I aim to answer is whether we have any reason to believe that the first-order realizer properties or the second-order properties are constituted by the fact that we have the capacity to respond to individuals as morally responsible and/or that we do fittingly respond in those ways. I argue that the traditional arguments for thinking that there is such a constitution relation between that capacity or the responses themselves and the conditions that render those responses fitting or appropriate fail.

I argue further that we can accept the correctness and supreme explanatory value of both a role-invoking analysis like (N) and a non-role-invoking analysis that only cites

---

<sup>4</sup> To be more precise, note that some accept a reasons-wise asymmetry here according to which the fact that S is blameworthy cannot itself be a reason for someone to blame S while the first-order realizers can be. I am not sure if I accept the reasons-wise asymmetry. I take it that S's being blameworthy can be a reason to blame S whether the individual blaming S is responding to the first-order properties that realize S's blameworthiness or the fact that S is blameworthy. On my view, if you respond to being told that Gena is blameworthy for negligently exposing her classroom to toxins only by *noting that* Gena is unreasonably careless in how she handles dangerous chemicals, then being apprised of the normative worth of her act did not serve as a reason to blame her. On my view, *just noting* a fact about Gena's behavior does not amount to blaming her or responding to her as responsible. Rather, as I discuss below, to blame Gena, you'd have to *recognize in* Gena's behavior some *ill will*. In other words, you'd have to *recognize that* her negligent behavior manifested a will that bears a quality not sufficiently good given the circumstances. So, if your response to being told about Gena's behavior involves a recognition of that sort and that recognition motivates certain blame-characteristic changes in your dispositional profile, then you blame her. Some might take this to imply that the fact that Gena is blameworthy did serve as a reason for you to blame her. For my purposes, it does not matter. I think I can remain neutral on whether we should adopt the reasons-wise asymmetry. What is important is that what is doing the normative, reason-giving work are facts about the individual's being worthy of the responsibility response which is grounded by first-order property realizers of a second-order property.

the properties that make moral-responsibility responses like praise and blame appropriate. So, if the properties F and G make an individual S *worthy of being* or *fit to be* the subject of a moral-responsibility response, the following analysis is no less correct:

(A) S is morally responsible (for  $\phi$ ) if and only if S has properties F and G.

A principle upshot of that first essay is that we can retain all of the important insights of a Strawsonian view even if we reject some key aspects of it. An aspect that is perhaps more widely thought to be the key strength of Strawson-style views is the famed

*Reactive Attitudes Account:* A moral-responsibility response involves a reactive attitude (i.e. a complex conative-affective attitude like resentment, indignation, guilt, gratitude, approbation, and pride).<sup>5</sup>

Indeed, many will think that some version of the Reactive Attitudes Account must be retained in order to offer an adequate response to the global skeptic. The idea is that alternative accounts incorrectly describe the phenomenon of responding to individuals as morally responsible and, so, the argument offered to justify that practice will miss its mark.

I disagree. In fact, I think my non-Reactive Attitudes Account is a central strength of my push against (P2), the Anti-Fitness Thesis because more accurately describes that phenomenon in a manner that preserves the elegance of the Reactive Attitudes Account. I call it, *the Recognition Account*, develop it in Chapter 3: *Recognition and Resentment*.

---

<sup>5</sup> Versions of the reactive attitudes account differ primarily in regard to which attitudes are admitted into the class of reactive attitudes that are constitutive of holding an individual morally responsible.

The key innovation of the Recognition Account is that it rejects the Humean idea that evaluative beliefs cannot motivate absent aid from independent desire. I develop the recognition account primarily by generalizing the Reactive Attitudes Account so that other non-conative-affective attitudes can also constitute our responsibility responses. The anti-Humean deconstruction of the reactive attitudes, then, generates the two-part specification of responsibility responses found in the

*Recognition Account:* A moral-responsibility response is constituted by an individual, *S*'s, recognition of an individual, *U*'s, quality of will where that recognition motivates a change in or reinforcement of a certain portion of *S*'s behavioral and attitudinal dispositional profile.

I show further that the anti-Humean moral psychology central to the recognition account is not susceptible to traditional objections to it and captures the contours of the phenomenon in question better than the Strawsonian account and George Sher's (2006) Humean account.

The main arguments that appear in the four essays that follow serve to defend and develop the Recognition Account while attempting to place it in a broader Strawsonian tradition. For instance, in the fourth essay (Chapter 4: *Restrictions and Resentment*), I defend the Recognition Account as an account that is promiscuous with respect to the sorts of attitudes that constitute responses like blame. Some worry that such accounts are objectionable in virtue of their promiscuity. Since it is sometimes unclear what the exact concern is, I consider three prominent ways to push the concern. I allay these worries and show that an account like mine that is less restrictive than versions of the Reactive Attitudes Accounts is no less viable. Hence, I not only defend the Recognition Account. I also place it in a broader, novel Strawsonian family.

In Chapter 5: *Self-Revelation and Resentment*, I continue to defend with an eye to locating the Recognition Account within broader Strawsonian scholarship. There I focus on concern related to the worry about promiscuity—namely, that views like the Recognition Account problematically civilize reactions like blame by rendering them superficial.<sup>6</sup> Others have directed this objection at accounts that adopt

*Self-Revelation:*       $\phi$  is attributable to an agent S as the basis of normative appraisal if and only if  $\phi$  reveals who S is as a practical agent.<sup>7</sup>

The main contributions of the fifth essay focus on developing Angela M. Smith's Self-Revelation View. First, I bring to the surface an under appreciated and under developed thesis of Smith's to show that Self-Revelation views are in fact deep. This also serves to unify accounts of moral responsibility precisely where many opt to adhere to non-unified accounts. But, this unity threatens to problematically flatten moral appraisals into monotone comments on and responses to practical agency. Second, then I show how a unified Self-Revelation Account can maintain a account of moral appraisals that does justice to their diversity.

I continue to develop the Recognition Account while placing it within a broader Strawson-inspired framework in Chapter 6: *Negligence and Resentment* by defending my account as a Quality of Will View. A key objection to Quality of Will Views is that they cannot capture the intuition that negligent injurers are blameworthy without positing that the negligent injurer harbored some bad attitude toward the injured or persons generally.

---

<sup>6</sup> This is Wolf's (1990) classic concern with real-self views.

<sup>7</sup> See, e.g., Waston (1996) and Smith (2008). Watson and Wolf refer to accounts that adopt something like self-revelation as *real-self views* and Smith refers to them as *self-disclosure views*. I refer to them as *self-revelation accounts* to avoid various things that the other names seem to imply.

Here, I show how an absent attitude or gap in reason can contribute to the quality of an individual's will just as much as the presence of an attitude can. Further, on my account, if an individual is capable of manifesting their will, they are fit to be held responsible. In response to a final concern, I appeal to the nature of our responsibility responses and argue that Strawsonians like myself are under pressure to think that failure to possess the capacity to understand responses like moral blame does not exempt an individual from the community of those fit to be blamed. Though my appeal and conclusion works as a response to an objection that grows from concerns about negligence, I am also attempting to show that the community of moral responsible agents on Strawsonian accounts is quite wide.

I continue to tow that line in Chapter 7: *Optimism and Resentment* and begin a more explicit push against the Global Skeptic. There I argue for and develop the view that the grounds that legitimate responding to an individual as morally responsible for what they've done remain in place for individuals who always experience everyone they encounter as natural and mechanistic things—i.e., those who have universally adopted the objective-stance. This view goes directly against the heart of all non-skeptical accounts of moral responsibility that are strongly inspired by Strawson as well as many of the views promoted by proponents of global skepticism.<sup>8</sup> I show that we can accept what I call *objective-stance optimism*. I argue further that with it we can offer an improved Strawsonian foundation in support of the view that our moral-responsibility concepts and responses remain coherent and justified whether causal determinism is true or not.

I follow up in Chapter 8: *Conclusion: Skepticism and Resentment* by putting

---

<sup>8</sup> See, e.g., Strawson (1962, 79), Wallace (1994, 23 – 40), Pereboom (2001, Chpt. 7; 2009; 2014, Chpts. 6 & 8), Sher (2006, Chpt. 1), and Sommers (2007).

forward a new version of Strawson’s argument that works to undermine support for (P2) of The Skeptical Argument—i.e., the premise that no one is fit (or deserves) to be responded to as a morally responsible agent. The version I offer uses insights picked up throughout my development of the Recognition Account. It also involves a further insight garnered from Elizabeth Lane Beardsley’s basically neglected essay, “Determinism and Moral Perspectives”, which preempts and in many ways anticipates the view put forward in Strawson’s highly influential, “Freedom and Resentment”.

To see how Beardsley figures into my response to The Skeptical Argument we need to briefly elaborate Strawson’s original argument. In “Freedom and Resentment”, Strawson attempts to reconcile (a) the pessimistic view that our moral responsibility concepts and practices are incoherent and unjustified in the face of causal determinism with (b) the optimistic view that we retain a sort of consequentialist coherency and justification of them despite the truth of causal determinism. The reconciliation is meant to show that our moral responsibility practices and concepts survive the truth of causal determinism without the need to interpret them simply in terms of mechanisms employed only to control and manage each other toward a common good. One main line against Strawson criticizes his reluctance to really raise and answer

*The External Question:* Whether we are psychologically incapable of ridding ourselves of the practice of holding each other and ourselves responsible or not, is anyone the fit subject of moral-responsibility responses, (i.e. is the practice of responding to individuals as normatively responsible for what they do externally justified)?<sup>9</sup>

---

<sup>9</sup> See, e.g., Wallace (1994, 95 – 109) and essays in McKenna & Russell (2008).

Strawson comes close to raising the question but ultimately and to the dissatisfaction of many has this to say about it:

...if we could imagine what we cannot have, *viz.* a choice in this matter, then 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. (Strawson, 1962, 83, his emphasis)

R. Jay Wallace succinctly characterizes the dissatisfaction many have with Strawson's position stating that he wrongly assumes that "the only possible reasons that might be cited for or against our practice of holding people morally responsible are pragmatic reasons, concerning the gains and losses of the practice for human life" (Wallace, 1994, 100).

It is in an attempt to supply a more satisfying answer to The External Question that I appeal to Beardsley.<sup>10</sup> First, however, I analyze Paul Russell's (1992) critique of Strawson and Wallace's (1994) critique of Strawson. Though I reject their final dismissal of Strawson, I draw from them the following desiderata that any adequate Strawson-style argument against The Skeptical Argument must meet:

- 1: It must support the view that we have no good reason to believe that determinism or lack of ultimate responsibility undermines one's ability to decide on the basis of reasons of mutual respect and regard;
- 2: It must support the assumption that the only way to determine whether we should abandon the concepts and practices centrally involved in moral responsibility responses is to appeal to the gains or losses such abandonment would bring to our human lives; and

---

<sup>10</sup> Boxer (2013) has also proposed a new version of Strawson's original argument that she thinks is promising. Though her discussion is insightful and careful, I critique and reject it before building my own.

- 3: It must contain something plausible to bridge the gap between (a) the conclusion that the truth of determinism does not bear on the coherency and appropriateness of our moral responsibility concepts and practices and (b) the notion that the only version of the External Question left to answer is one where the propriety it concerns is (in large part) answered by considering the gains and losses of continued deployment of moral responsibility concepts and practices.

After I critique and gain further insight from K. E. Boxer's (2013) unsuccessful reconstruction of a Strawson-style argument, I appeal to theses and arguments developed in the foregoing chapters to accomplish (1). I follow by relying on Beardsley to meet (2) and (3).

In her essay, Beardsley identifies three moral dimensions. Two are the primary dimensions against which we make moral appraisals. Roughly, one is the dimension we occupy when appraising character and the other dimension is taken up as we appraise the extent to which an individual's decision to  $\phi$  meets a certain moral standard. The remaining moral dimension is the one we occupy when wonder whether an individual is *ultimately* responsible for their character or their decision to  $\phi$ .

According to Beardsley, this last question is only intelligible due to (I) the fact that questions about an individual's moral agency in the first two dimensions are intelligible, (II) the fact that the appraisals made while occupying a position in either of the first two moral dimensions are appropriate, and (III) the fact that some of those appropriate appraisals offer a *normative* evaluation of an individual's character as good or bad and/or recognize an individual's decision to  $\phi$  as morally praiseworthy or blameworthy. Beardsley concludes, then, that there may be some sense to the claim that we are not ultimately the originators of our characters or our decisions, and that sense is

parasitic on the fact that we are nevertheless fit subjects of moral or normative appraisal in the first two dimensions.

With this background I aim to meet desiderata (1) – (3). First, I argue that we have good reason to believe that determinism undermines our status as morally responsible agents *only if* (i) we must give a negative answer to the question, “Is anyone ever ultimately responsible for doing what they did?” and (ii) that negative answer entails that no one is fit for moral appraisal.<sup>11</sup> Second, by appealing to my foregoing development of the Recognition Account, I show that neither the truth of determinism nor our lack of ultimate responsibility gives us (a) a reason to believe that individuals cannot manifest their will, or (b) a reason to believe that their will cannot be of good or poor quality. The upshot is that we can affirm the truth of determinism *and maintain* that we might nevertheless be fit for moral appraisal. Third, I argue similarly that can acknowledge that we lack ultimate responsibility and maintain our fitness for praise and blame. Thus, desiderata (1) is met—neither determinism nor lack of ultimate responsibility entails a lack of moral responsibility.

I thereby establish the claim that from the perspective of moral responsibility, what matters is whether we have the capacity encounter, recognize, and respond to reasons in a normatively significant manner. Further, lacking that capacity, which is the only thing that should lead us to believe the Anti-Fitness Thesis from the Skeptical Argument, would be abnormal. And, if the thesis of causal determinism or ultimate responsibility entails that such a failure is normal and universal, then both are muddled

---

<sup>11</sup> Cf. Boxer (2013). Boxer deals with the issue almost exactly as I do but appeals to H. L. A. Hart’s views about punishment to reject the ultimate responsibility concern. I critique her approach in my concluding chapter.

and incoherent. Thus, we have no reason to believe the Anti-Fitness Thesis.

Furthermore, as Beardsley suggests, our negative answer to the question “Is anyone ever ultimately responsible for doing what they did?” does not entail that no one is fit for moral appraisal.<sup>12</sup> So, as Strawson says we can wonder what we gain and what we lose from participation in and perpetuation of the practice of responding to each other as morally responsible and that may result in the conclusion that we should continue the practice or take aims to change or eliminate it. But, the truth of determinism does not bear on this consideration. So, desiderata (2) and (3) are also met.

An important fact of our moral lives is that we respond to individuals as normatively responsible. When the crowded market place is bombed, where bursting financial bubbles spur widespread economic collapse, when gun smoke is alive in an elementary school, we hold responsible and we blame. At a friend’s loss, at a friend’s gain, when children and adults learn despite the odds we respond to individuals as normatively responsible, we blame, and we praise. When a sea change comes supporting debt forgiveness for those struggling in a null-growth economy, when peace emerges between enemies, when a cure is found, we find ourselves facing situations that simply call out for a response. In response, we recognize and appraise for ourselves and each other. A world devoid of this phenomenon would be extremely unfamiliar. For better or worse, responsibility responses are central to our moral experience. The controversial question is not *whether* we respond to individuals as morally responsible agents but whether anyone actually *is* morally responsible. My aim here is to lay the foundation for answering this question affirmatively. I turn now to that task.

---

<sup>12</sup> Arpaly (2006) is a key influence for the argumentative strategy that I adopt here.

## II:    NORMATIVISM AND RESENTMENT†

Central to a fitting-attitudes account of something,  $y$ , is an analysis of the form<sup>13</sup>

*FA*:     $x$  is  $y$  just in case it is fitting/appropriate to have attitude  $A$  toward  $x$ .

So, for example, a fitting-attitude account of value at its core includes something like

*FA-Value*:     $x$  is valuable just in case it is fitting/appropriate to have an attitudinal value-response toward  $x$ .

And, many employ the thought that the responses to an individual's *being* morally responsible should be central to an explanation of her being so in a view, often referred to as *normativism*, at the center of which is

(N)     $S$  is morally responsible (for  $\phi$ ) if and only if it is appropriate to respond to  $S$  as morally responsible (for  $\phi$ ).<sup>14</sup>

Some normativists also endorse

*RD*:    The properties that render responsibility responses fitting are *response-dependent* insofar as *being worthy of responsibility responses* depends on the capacity to respond to individuals as morally responsible.<sup>15</sup>

One of the main things I do here is argue that *RD-Normativism* is not supported by the arguments advanced by proponents of it.

---

<sup>13</sup> For an overview of such accounts, see, Rabinowicz & Ronnow-Rasmussen (2004, 394 – 400).

<sup>14</sup> Slightly modified from Wallace (1994, 91). Cf. Beardsley (1960), Strawson (1962), Dennett (1984, Chpt. 7, esp. 158 – 172) and Fischer & Ravizza (1998, Chpt. 1, esp. 5 – 8 and 8 n. 12), Pettit (2001), Helm (2011), and McKenna (2012, esp. 34 – 55).

<sup>15</sup> See, e.g., Helm (2011), McKenna (2012, 34 – 55), and Shoemaker (forthcoming (b)).

Normativism arose primarily from attempts to advance a picture of moral responsibility aligned with P. F. Strawson's on which the interpersonal nature of responsibility practices is indispensable to a correct explanation of moral responsibility.<sup>16</sup> Clearly, (N) offers a way to capture this thought. And, RD-Normativism allows this perspective to seep through to the conditions that make responding to an individual as morally responsible fitting. Indeed, the arguments RD-Normativists use to support their view lean on the indispensability-thought which is in turn taken to nicely capture important themes from Strawson. My push away from RD-Normativism, then, involves separating it from the structure suggested by the indispensability-thought and certain Strawsonian themes. To do so, I argue that (a) the structure does not favor RD-Normativism (§2), nor (b) does adherence to important Strawsonian themes require it (§3). First, I delineate a problem facing normativism in order to flesh out its contours, further distinguish it from RD-Normativism, and suggest a solution of my own that does not appeal to RD (§1). Given the theoretical neutrality amongst normativists, my way of handling the problem which only employs a distinction in types of explanation is preferable to a solution that requires the adoption of RD. The broader implication of the chapter, then, is that we can and should maintain the outlook that the interpersonal nature of responding to each other and ourselves as morally responsible is indispensable to explaining moral responsibility even if RD-Normativism is unsupported and unmotivated. I turn, then, to exploring normativism in light of *the middle-man problem*.

---

<sup>16</sup> See, e.g. Scanlon (1981; 2008, Chpt. 4), Dennett (1984, Chpt. 7), Wolf (1991, Chpt. 1, esp. 15 – 22, and Chpt. 4, esp. 70 – 73), Wallace (1994; 2011), Korsgaard (1996, Chpt. 7, esp. 197 – 199), Oshana (1997; 2004), Fischer & Ravizza (1998, Chpt. 1, esp. 5 – 8 and 8 n. 12), Pettit (2001), Darwall (2006, Chpt. 3), Helm (2011), Shoemaker (2011), McKenna (2012). Cf. Midgley (1984, Chpt. 3; 1994, 104 – 105).

## 1: THE MIDDLE-MAN PROBLEM

The middle-man problem arises from the fact that normativist adopt (N) and posit views about what makes moral-responsibility responses appropriate or fitting like

- (C) S is morally responsible (for  $\phi$ ) if and only if A where A is the set of properties one must possess in order to be appropriately responded to as morally responsible (for  $\phi$ ).<sup>17</sup>

The middle-man, as it were, is (N). The general puzzlement, then, is that it is unclear why normativists don't simply accept (C) as the more explanatorily fundamental or only precisely correct analysis of being morally responsible.

For instance, on Wallace's view, appropriateness is filled out in terms of fairness. According to Wallace, it is fair to hold one responsible just in case no excusing or exempting conditions are applicable to the one being so held.<sup>18</sup> Wallace argues that no excusing or exempting conditions are applicable where the action was performed by an agent who possesses the power of reflective self-control, which is the "rational power to grasp and apply moral reasons, and to control their behavior by the light of those reasons" (Wallace, 1994, 1).<sup>19</sup> Thus, as Manuel Vargas (2004, 224) points out, in addition to (N),

Wallace is committed to:

- FAIR: S is morally responsible (for  $\phi$ ) if and only if it is fair to hold S responsible (for  $\phi$ ).

---

<sup>17</sup> Cf. Vargas (2004), Arpaly (2006, 6 and 28), and Fischer (2008, 137 – 142).

<sup>18</sup> Wallace (1994, Chpt. 4, esp. 103 – 117, Chpts. 5 & 6).

<sup>19</sup> The development of Wallace's account occurs throughout the book, but surfaces more and in Wallace (1994, Chpts. 4, 5, & 6).

So, it seems Wallace should accept something like<sup>20</sup>:

(C-W) S is morally responsible (for  $\phi$ ) if and only if S possesses the power of reflective self-control.

Indeed, since (C-W) identifies that in virtue of which one is morally responsible, it seems to be at the very least no less accurate and no less fundamental as an explanation than (N). But, as a normativist, Wallace denies that such is the case.<sup>21</sup> So, the middle-man problem pushes us to ask Wallace: why not accept a restatement of (N) that jettisons reference to responses?

Bennett Helm (2011) offers a solution to the middle-man problem which adopts Wallace's (C-W), RD-Normativism, and

*Constitutionalism*: The source of reasons for responding to an individual as morally responsible (for  $\phi$ ) are constituted by a pattern of fitting exercises of her capacity to respond to individuals as morally responsible (for  $\phi$ ).

For Helm, the capacity to respond to individuals as morally responsible is the capacity for rational self-control and the source of reasons that is constituted by a pattern of fitting exercises of that capacity is dignity. The idea, then, is that (C) is rejected as a correct analysis of being morally responsible (for  $\phi$ ) due to the fact that RD and Constitutionalism are true but (C) ignores or at least hides those facets of the analysandum.

---

<sup>20</sup> Wallace (1994, 2) summarizes the rational power to grasp and apply moral reasons, and to control their behavior by the light of those reasons as the power of reflective self-control.

<sup>21</sup> Fischer (2008, 137 – 142).

More specifically, central to Helm's solution is the view that "rational self-control itself is intelligible only in terms of one's being held responsible" (Helm, 2011, 232). According to Helm, the only rational self-control relevant to moral responsibility is a kind that "differs from [the] self-control a dog exhibits in delaying snatching the treat balanced on its nose in that it is responsive to the *dignity* of others—in that, we might say in a Kantian vein, it is motivated out of duty" (Helm, 2011, 233, my emphasis). According to Helm, an individual's dignity is itself constituted by rational patterns of responding to her as responsible, where such instances are comprised of having one of the reactive attitudes (e.g. resentment, indignation, or guilt). From this, Helm gathers that

The upshot is that whereas Wallace claims that being responsible requires being appropriately held responsible which in turn requires that one can exercise rational self-control [...], we can now see that exercising the relevant sort of rational self-control [...] require[s] that one be a fellow person with others in a community of respect in which each can hold the others responsible for their actions. (Helm, 2011, 233)

So, on Helm's view, normativists should retain something like the middle-man (N) and reject (C), since RD and Constitutinalism imply that being morally responsible depends on and is constituted by a certain collection of fitting responsibility responses.

At least two problems plague Helm's solution. First, nothing Helm says supports the view that possessing a capacity like rational self-control requires exercising it, fittingly or not. And, furthermore, it seems that one can have the capacity without responding to reasons of any sort as well as have it in the absence of such reasons. That is, even if there were reasons grounded by dignity that supported holding an individual responsible, it seems that one could fail to respond to those reasons and still possess the capacity of rational self-control. Perhaps she swallowed a pill that perfectly blunts her

receptivity to reasons to respond to individuals as responsible while retaining the ability to control herself on the basis of reasons. Alternatively, even if one lived in a world where no one held anyone responsible and, so, no dignity was constituted, one could still have the capacity to respond to the reasons grounded by dignity were they to arise.<sup>22</sup> Thus, at the very least, Helm either needs to adjust his devotion to (C-W) or show that we are wrong to think that the capacity could be retained in those circumstances.

Second, unless we adopt a constitutionalist story about the source of at least one important type of moral reason, we cannot accept Helm's solution.<sup>23</sup> To see this, note that rational self-control by Helm's own lights is the ability to act on and for reasons grounded in or flowing from our dignity. But, for Helm dignity is constituted by our having rational patterns of reactive attitudes which centrally involves the exercise of the capacity of rational self-control. Surely, there is a viable normativism that leaves open whether dignity as a source of moral reasons is constituted by that for which it provides reasons. Such a substantive thesis may be correct, but it is a commitment that cannot be adopted without serious theoretical baggage. Thus, normativists not already committed to a substantive thesis like constitutionalism about the source of reasons for responsibility responses should look elsewhere for a solution.

Perhaps adopting the view that *being worthy of moral-responsibility responses* is a response-dependent property offers a more amenable alternative for answering the middle-man problem. At the very least this alternative seems to avoid adherence to a

---

<sup>22</sup> A closely related problem stems from the fact that, at best, Helm has provided an analysis of what it takes to be morally responsible for one's conduct. But, Wallace and others are equally concerned to show that holding responsible (i.e. the reactive attitudes) must be featured in the analysis of being a morally responsible agent not just the analysis of being morally responsible for conduct. For useful discussion here, see, McKenna (2012, 34 – 39).

<sup>23</sup> Interestingly, Wallace would seemingly not accept Helm's solution. See, e.g., Wallace (2004).

Helm's Constitutionalism. Furthermore, this alternative might sit well with those normativists who hope to retain a view friendly to naturalism that captures the intuition that being morally responsible (for  $\phi$ ) is not mind-dependent. Of course, it only captures that intuition by claiming that being morally responsible (for  $\phi$ ) is independent of any particular individual's mind, since even on the response-dependent thesis, RD, *being worthy of moral-responsibility responses* still depends on our having a *mental* capacity. Nevertheless, it is an alternative that I explore in the next section. There I will argue that no RD-Normativist has offered a viable argument for thinking that being morally responsible is response-dependent.

Before moving on, I want to be clear that, as a normativist, I face the middle-man problem as well. But, my solution does not appeal to response-dependence. Rather, it employs the distinction between a role-invoking explanation and a non-role-invoking explanation. A role-invoking explanation cites the role something plays in our lives. A non-role-invoking explanation does not. And, since it is a pragmatic matter whether the role-invoking explanation should be used in an attempt to explain something like moral responsibility or not, one type of explanation may be better for explaining some things while the other type is prized elsewhere. My solution, then, is simple. Normativists should accept (N) as a correct analysis that is deployable in situations calling for a role-invoking explanation and they should adopt the relevant (C) analysis as a non-role-invoking version of (N) that is no less true and no less fundamental than (N). The neutrality of my solution as compared to the constitutionalist solution or a solution involving a commitment to response-dependence about certain normative properties speaks to its superiority unless we have good reason to think that the response-

dependence thesis is true. I turn, then, to see what RD-Normativists have said in favor of that view.

## 2: SUPPLEMENTATION, DERIVATION, AND DISTINCTION

Traditionally, support for (N) comes from two directions. Each direction grows from an attempt to develop the intriguing idea that focusing on the phenomenon of holding an individual responsible can (a) help elucidate moral responsibility and (b) show that our practice of holding responsible is not threatened by the truth of causal determinism. The need for development arises primarily from the thought that Strawson's attempt to save moral responsibility via a picture that places a certain conception of the phenomenon of responding to an individual as responsible does not actually stave off the threat posed by the truth of causal determinism. The concern is that all Strawson's picture does in this regard is show that we have an unshakable commitment to the practice.<sup>24</sup>

What is typically sought after, then, is some way of *justifying* our practice of treating ourselves and each other as morally responsible agents in the face of causal determinism. Wallacian normativists argue that we should adopt (N) and only (N) for this task. So far, normativists have offered two reasons for their position. On the one hand, they argue that (N) clearly delimits a path that we can take that will allow us to justify responsibility responses in a manner that stays true to the Strawson's framework without taking on board other worrisome meta-ethical theses or obscuring the moral

---

<sup>24</sup> See, e.g. Russell (1992); Wallace (1994, Chpt. 4, esp. 95 – 103); Magill (1997). I do not mean to imply that the authors cited are convinced that we must adopt (N) or something similar in order to develop Strawson's thought in a manner that will secure it against the threat of determinism. For instance, Magill defends Strawson's naturalism and develops Strawson's view along naturalist lines without (N).

responsibility debate. On the other hand, RD-Normativists claim that (N) is (1) no more metaphysically robust than it needs to be and (2) it is metaphysically weak enough to be comfortably absorbed by the naturalist.<sup>25</sup> Analysis of each traditional line of argument shows (i) that neither supports RD-Normativism, (ii) that neither supports adopting (N), and (iii) that neither supports adopting (N) as the only correct analysis of being morally responsible. I begin my analysis with the first route.

### 2.1: The First Route and The Supplementation Thesis

Both of the traditional routes taken to support (N) are best understood against a background of the accounts of responding to an individual as responsible that are accepted by many normativists.<sup>26</sup> Each account belongs to a family of views on which holding responsible involves one or more conative-affective attitudes. I call them, *reactive attitudes accounts*.<sup>27</sup> Different versions of reactive attitudes accounts can be generated by identifying a different set of reactive attitudes as those involved in responsibility responses. For example, Strawson, Wallace, Michael McKenna, take reactive attitudes to be sentiments like resentment, indignation, guilt, gratitude, approbation, and pride. But, Wallace and McKenna hold that only resentment, indignation, and guilt count as responsibility responses.<sup>28</sup> Whether one adopts a reactive attitudes account and whether one adopts a wide or narrow version is not important for

---

<sup>25</sup> See, e.g., Wallace (1994, Chpt. 4, esp. 85 – 95), and McKenna (2012, Chpt. 2, esp. 39 – 45).

<sup>26</sup> Of course, normativism does not entail a particular account of holding responsible. It is just a sociological fact that I am stating. Further, many non-normativists also adopt these sorts of accounts of holding responsible.

<sup>27</sup> McKenna (2012, 45) calls them conative-affect accounts.

<sup>28</sup> Wallace (1994, 25 – 40; 2011). Cf. McKenna (2012, 27).

my purposes. Rather, the important point here is that a worry arises for those that (a) accept the Strawsonian view that the interpersonal nature of our practice of responding to individuals as responsible is indispensable to explaining moral responsibility (i.e. the indispensability-thought) and (b) accept a reactive attitudes account of that practice. The worry is that the combination of (a) and (b) entails an expressivist account of being morally responsible.<sup>29</sup>

Support for the idea that (a) and (b) lead to expressivism about moral responsibility is that (a) claims that the analysis of being morally responsible must somehow include the essentially interpersonal phenomenon of responding to an individual as responsible. And, (b) claims that that phenomenon is constituted by attitudes with content that can be neither true nor false. Hence, the idea goes on, accepting (a) and (b) entails that when an individual U claims that another individual, S, is morally responsible and blameworthy for behaving in some way,  $\phi$ , U merely expresses that she has some sort of negative, attitude toward S's behavior where that attitude is not truth-evaluable. The first route to supporting the view that (N) is the only correct analysis of being morally responsible begins here.

Basically, on the first route, the idea is to lean on the plausibility of accepting (a) and (b) and the desire not to end up with the conclusion that seems to follow—namely, the conclusion that it is neither true nor false that individuals are morally responsible. So, more precisely, the thought is that (a) and (b) are appealing, but a suitable understanding of being morally responsible will at a minimum not commit one either “to the noncognitivist position that claims about moral responsibility lack truth-conditions, or to

---

<sup>29</sup> See, e.g., Bennett (1980). Also, Laura Wadell Ekstrom (2000, 146 – 149) describes Strawson's view as an expressivist account of moral responsibility.

the eliminativist position that there are no facts about moral responsibility” (Wallace, 1994, 90). The first route unfolds, then, as a search for how to nuance acceptance of (a) and (b) in order to avoid those commitments. Wallace offers the most filled out version of this line of thought.

After developing his version of a reactive attitudes account, Wallace considers and rejects

- (D) S is morally responsible (for  $\phi$ ) if and only if we are disposed, under favorable conditions, to hold S morally responsible (for  $\phi$ ).

Wallace is concerned that (D) does not provide “the best framework for approaching questions about the conditions of responsibility” primarily for the reason that “it does not allow us to formulate perspicuously the incompatibilist position” (Wallace, 1994, 90).<sup>30</sup>

Explaining, he writes:

To see this, note that if determinism is true, and incompatibilism correct, then nobody ever is or was morally responsible for what they do. And yet for all that, we might still be saying that the incompatibilist is making a normative claim—the claim, namely, that it would be inappropriate or incorrect to hold people responsible if determinism is true. This is quite different from saying that we would be disposed not to hold people responsible if determinism is true. The incompatibilist maintains that whatever might be the case concerning our dispositions to hold people morally responsible, determinism would make it inappropriate or incorrect to acquiesce in that practice, giving us a reason to refrain from holding people responsible that applies across the board. (Wallace, 1994, 90 – 91, two footnotes omitted)

Wallace admits in a footnote that one could, following Michael Smith (1989), allow that

---

<sup>30</sup> John Martin Fischer & Mark Ravizza (1993, 18) and McKenna (2012, 34 – 35) follow Wallace in rejecting (D) for its lack of explicating the normative conditions involved in holding one morally responsible.

the ‘favorable’ conditions are the “conditions of ideal rationality” so that “what the analysis is interested in is not really whether we are disposed to respond in a specified way in given circumstances, but whether it would be rational or appropriate to do so” (Wallace, 1994, 90 n. 11). Nevertheless, he holds that his proposal, (N), is “more perspicuous [...] in bringing normative questions explicitly to the foreground” (Wallace, 1994, 90 n. 11). Thus, Wallace supplants (D) with (N), and his reasons for doing so comprise the first route to supporting (N) as the only correct analysis of being morally responsible.

To be clear, the basic idea is that, first and foremost (N) is commensurate with acceptance of (a) the indispensability-thought and (b) a reactive attitudes account. Furthermore, Wallace’s idea is that we can escape a commitment to non-cognitivism and/or eliminativism, given that adopting (N) commits one to the view that the attitudes involved in responding to an individual as responsible can be evaluated for appropriateness. And, further still, we avoid these commitments in a manner that succinctly and correctly frames the dispute between incompatibilists and compatibilists. Given this, Wallace maintains that those of us who adopt (a) and (b) and hope to escape non-cognitivism/eliminativism about moral responsibility in a manner that best frames the debate should accept (N) as the only correct analysis of being morally responsible.

While insightful in many ways, the first route fails as a reason in favor of adopting (N) and no other analysis. To see this, notice first that we are given no explanation yet as to why *being* morally responsible is characterized in terms of *holding* responsible.<sup>31</sup> Presumably, since the objection to (D) focuses on its failure to

---

<sup>31</sup> This is a result of my set-up and not a criticism of Wallace. Below I attend to reasons for characterizing

perspicuously formulate the incompatibilist position, a better analysis will be *any* analysis that cites more than our various dispositions. The inability of (D) to neatly formulate the debate is taken to be a result of the fact that it only cites our disposition to hold individuals responsible. The failure is taken to persist even if we supplement (D) with a claim about the appropriateness of manifesting those dispositions. According to Wallace, we need the normative in there as well.

That short line of reasoning about (D) in favor of (N) is not very convincing as a reason to accept the view that (N) is *the only* correct analysis of being morally responsible. Notice that we can agree that analyses like (D) are insufficient for a full story about being morally responsible for the reasons Wallace cites—namely, we need an analysis that is role-invoking and normative.<sup>32</sup> Be that as it may, Wallace offers us no reason to reject a suitable version of (C) as a non-role-invoking analysis of being responsible that is just differently informative but *no less accurate* than (N).

Thus, all that the first route to the view that (N) is the only correct analysis of being morally responsible licenses is

*The Supplementation Thesis:* We must supplement *non-role-invoking* accounts of being morally responsible with a suitable *role-*

---

being responsible in terms of holding responsible. As a precursor, roughly, it is a reduction of sorts.

<sup>32</sup> Wallace might be offering an open question argument against the adoption of non-role-invoking analyses like (D) and versions of (C). For instance, he might be saying that we can still meaningfully ask, “Is being morally responsible (for  $\phi$ ) to be disposed under the conditions  $C$ ?” for any set of conditions, and, so, (D) must be incorrect. If that’s right, then it would at least be on target insofar as it might stand a chance of giving us a good reason to reject such analyses. However, open question arguments are simply bad arguments. Of course, defense of that claim might require more than a footnote. Suffice it to say, however, that a strong reason to reject open question arguments is that there are many things that we identify with other things where we now know that the identification holds *and* asking whether it does still seems meaningful. For example, we didn’t know but now we do know that water is roughly the chemical compound of two hydrogen atoms and one oxygen atom—i.e.  $H_2O$ . But, the question, “Is water  $H_2O$ ?” seems like a meaningful and open question in many respects.

*invoking, normative* account concerning the appropriateness of the attitudes involved in responsibility responses.

Clearly, the supplementation thesis and (N) are not antithetical. But, just as clearly, they are not identical. Thus, the first route fails to support *exclusively* going in for (N), for it makes it no less reasonable to accept (N) *and* some suitable analysis where only the conditions that would make responding to an individual as responsible appropriate are featured (like (C) above) rather than any reference to the phenomenon itself. So long as we also adopt a role-invoking account of *being worthy of responsibility responses* commensurate with that version of (C) we do not run afoul of the first route. And, nothing Wallace says supports the view that being morally responsible is response-dependent. Thus, even if we accept Wallace's claims, we need not accept that (N) is the only correct analysis of the target property or RD-Normativism. The second route fares no better.

## 2.2: The Second Route and The Derivation Thesis

In support of exclusively adopting (N) as the correct analysis of being morally responsible, Wallace identifies two ways to understand the facts about being morally responsible. One way to understand these facts can be identified as *the extreme metaphysical interpretation*.<sup>33</sup> On this interpretation, "we would suppose that there is a

---

<sup>33</sup> McKenna (2012, 42 – 43). McKenna is basically following Wallace (1994, 87), and I am following McKenna in thinking that Wallace's terminology does not parse the positions aptly. It is unclear to me why we should consider this non-reductive account as an extreme metaphysical position and a reductive or dependence account as less extreme in this sense. I take it that rhetoric may play a significant role in the

fact of the matter about responsibility “in itself,” a fact about what it is to be *genuinely* or *really* responsible, and that this fact is prior to and independent of our practice of treating people as morally responsible agents” (Wallace, 1994, 87, his emphasis). An alternative way to understand the facts about being morally responsible can be identified as *the moderate metaphysical interpretation*. On the moderate metaphysical interpretation, facts about being morally responsible are in some way dependent on facts about responding to an individual as responsible. Wallace opts for a moderate metaphysical interpretation because “it seems incredible to suppose that there is a prior and independent realm of facts about responsibility to which such emotions [i.e. reactive attitudes] and actions should have to answer” (Wallace, 1994, 88).

Wallace offers the dispositional analysis, (D), seen above, as a first attempt to locate an account friendly to the moderate interpretation. In other words, he offers (D) as a first attempt at an analysis of being morally responsible that is more palatable than those analyses that count as extreme metaphysical interpretations. The offering is an example of taking the second route to supporting RD-Normativism and/or the view that (N) is the only correct analysis of being morally responsible. Rejecting (D) for the reasons we visited in the previous sub-section, Wallace accepts (N) and in so doing offers it as a palatable alternative analysis. This move comprises Wallace’s version of the second route. However, accepting either RD-Normativism or the view that (N) is the only correct analysis on this basis is far too quick.

To see this, note that the second route may help to explain why being morally responsible is characterized in terms of moral-responsibility response in both (D) and

---

choice of terminology by the two authors. Here, I adopt the terminology to simply reduce confusion between the texts and my argument.

(N). But, the fact, if it is a fact, that we should resist the sorts of interpretations of moral responsibility that count as extreme metaphysical interpretations does not on its own support exclusively adopting (N). For one, (N) seems not to have risen from the fact that we participate in responsibility practices. Thus, it seems Wallace should find it no less surprising that our emotions and actions answer to (N), which is itself part of that prior and independent realm of facts about responsibility.<sup>34</sup>

Moreover, the second route at most gives us a reason to accept *any* account that derives what it takes to be morally responsible from our best theory about responding to individuals as responsible. And, taking insights from the reasons supporting The Supplementation Thesis, we might hold that the second route gives us reason to adopt

*The Derivation Thesis:* We should only accept those accounts of moral responsibility that derive what it takes to be morally responsible from the conditions under which holding responsible is appropriate.

But, like the supplementation thesis, The Derivation Thesis champions (N) no more than it champions analyses where facts about the properties an individual must possess in order to be appropriately held responsible replace the analysans in (N). For instance, an analysis like (C-W) that cites rational self-control as necessary falls in line with the dictates of The Derivation Thesis. This is simply due to the fact that the types of analyses that will fit the bill issued by The Derivation Thesis outstrip (N) and includes analyses like (C-W). To see this, we only need to see that we could have come to adopt (C-W), for example, by observing that it is appropriate to respond to an individual, S, as responsible (for  $\phi$ ) only where they had rational control (over  $\phi$ -ing). Thus, even if we

---

<sup>34</sup> I thank Mark van Roojen for helping me see this important point.

adopt The Supplementation Thesis and The Derivation Thesis, we only have a reason to pair an analysis like (C-W) or (D) with (N) in our account of moral responsibility.

Still, McKenna might worry that my critique overlooks an important distinction and is for that reason too quick. In the next and final sub-section, I delineate a response available to proponents of Wallacian normativism in light of novel distinctions put forward by McKenna in his recent attempt to support that view. I argue that the response fails.

### 2.3: McKenna's Distinctions

McKenna makes at least two important contributive distinctions. The first is a disambiguation of (N) that appears to be implicit in the original statement of it. The second is a way of dividing up accounts of moral responsibility that may be taken to be more exact than the division between extreme metaphysical interpretations and moderate metaphysical interpretations. Although the disambiguation and distinctions are helpful, I show that they do nothing to improve support for the view that (N) is the only correct, most fundamental analysis of being morally responsible or RD-Normativism specifically. Consider the disambiguation first.

McKenna correctly points out that (N) is ambiguous due to its incorporation of the parenthetical phrase "(for  $\phi$ )". As McKenna notes, without the phrase, (N) appears to

speak to S's agency whereas with the parenthetical it speaks to items attributable as the basis of moral appraisal for conduct.<sup>35</sup> Elaborating, McKenna writes:

In the interest of plausibility, the appropriateness of holding morally responsible *for an action* as mentioned in (N) must be understood to offer a *pro tanto* reason—a legitimate reason albeit one that could be defeated by other weighty reasons—rather than an all-things-considered reason. Clearly there are times when a person is morally responsible for an action, perhaps, one for which she is blameworthy, and it would *not* be all-things-considered appropriate to blame her or hold her morally responsible at all. Maybe doing so would cause the destruction of the planet, or maybe, more realistically, it would be too much bother since the moral infraction is so small and one's personal circumstances are sufficiently burdensome that it would not be worth the cost. On the other hand, the reasons for thinking it appropriate to hold a person morally responsible in the sense of regarding her to have, and treating her as if she does have, the status of being a morally responsible agent are ones that should *not* be characterized as *pro tanto* but rather all things considered. (McKenna, 2012, 36 – 37, emphasis his, footnote omitted)

In light of this, McKenna separates (N) into the following<sup>36</sup>:

(N-Action): S is morally responsible for action x in the sense of being praiseworthy or blameworthy for x if and only if it would be *pro tanto* appropriate to hold S morally responsible for action x by praising or blaming her.

(N-Agency): S is a morally responsible agent if and only if, all-things-considered, it would be appropriate to hold S to be a morally responsible agent, one who is liable to praise and blame.

(N-Action) is interesting in its own right. But, focus for a moment on (N-Agency).

---

<sup>35</sup> I have already incorporated this useful distinction throughout this paper. It is McKenna's further insights that I'm interested in bringing to the reader's attention.

<sup>36</sup> McKenna (2012, 38, punctuation error corrected and grammatically clarified replacing 'it' with 'x').

On the one hand, it is unclear why someone like McKenna who renounces (D) should go in for (N-Agency), since it appears to smuggle dispositions back into the analysis.<sup>37</sup> Notice, to be liable to praise and blame as it appears in (N-Agency) is simply to be susceptible to being praised and blamed. It is natural to interpret ‘liability’ here as one’s being a potential trigger of the disposition to praise or blame. Thus, (N-Agency) seems to interpret being a morally responsible agent as being susceptible to being praised or blamed. Thus, it allows dispositions to resurface in the explanation of moral responsibility at least with respect to agency.<sup>38</sup>

Indeed, in explaining (N-Agency), McKenna makes it clear that the ‘hold to be’ is not just “a theoretical judgment about the agent” (McKenna, 2012, 38). On his view, ‘hold to be’ here “comes to more than the belief or judgment that the person is a morally responsible agent [...], it is also a practical judgment involving the regard one has for the agent, an attitude or stance one adopts toward the agent in which one is interpersonally prepared and *disposed to respond in certain ways*, were the agent to do the sort of thing that would make it (minimally, *pro tanto*) appropriate to praise or blame” (McKenna, 2012, 38, my emphasis).

Although difficult to parse, I want to glean two things from McKenna’s explanation that seem to shine through. One is that McKenna’s explication supports my suspicion that dispositions are being smuggled back into the explanation, which may be in some ways harmless given Smith (1989) as Wallace points out. The other is that

---

<sup>37</sup> McKenna (2012, 34 – 37) basically renounces (D) for the same reasons as Wallace.

<sup>38</sup> Note the comma separating ‘a morally responsible agent’ and ‘one who is liable to praise and blame’. Certainly not damning evidence, but very suggestive evidence in support of my claim about (N-Agency) and how it interprets being a morally responsible agent.

McKenna's (N-Agency) seems to conflate a *manifestation* of certain dispositions with two things—namely, those particular dispositions *and* that which structures them.

To see this, note that a natural way to understand the 'hold to be' in (N-Agency) is as a manifestation of the disposition *to respond to an individual as a morally responsible agent*. And, it is natural to take this sort of disposition to be structured by the stance alluded to in McKenna's explanation—namely, the *participant stance*. Coleen Macnamara offers an excellent description of that stance:

To adopt this broad psychological orientation toward another is to regard another as a responsible agent, where such regard is understood as more than a matter of simple belief. It is a broad orientation that guides our *presumptive interpretations* of the person's conduct—for example, we presume until presented with evidence to the contrary that the person is responsible for her conduct. It also structures our *behavioral dispositions*—we are disposed, for example, to offer and ask for reasons from the other. And it structures our *emotional proclivities*. (Macnamara, 2011, 84, her emphasis, footnote omitted)<sup>39</sup>

So, the 'hold to be' in (N-Agency) is ambiguous between responding to or interacting with a particular individual in a particular circumstance and a broad psychological orientation that carries across individuals and circumstances. Thus, McKenna's (N-Agency) is at the very least in need of repair and further elucidation. Furthermore, it should be clear that the disambiguation, while perhaps helpful in some ways, makes neither of the two routes canvassed above more supportive of (N) let alone more supportive of the view that we should exclusively adopt (N) or RD-Normativism.<sup>40</sup>

---

<sup>39</sup> This stance is hinted at in Beardsley (1960, 4 – 7 & 16 – 17). In a footnote omitted from the quoted passage, Macnamara directs us toward Maggie Little's (1998) article on various stances.

<sup>40</sup> In case this is not clear, consider what each of the two routes actually supports, if anything. The first route, if sound, supports The Supplementation Thesis, which is the thesis that we must supplement any *non-normative* account of being morally responsible with a *normative* account concerning the

Rather, it simply unpacks (N) in order to offer a more precise view.

Turn, then, to the distinction McKenna puts forward as an improvement upon Wallace's distinction between extreme metaphysical interpretations of moral responsibility and moderate metaphysical interpretations. Rather than appeal to Wallace's distinctions in his version of the second route to exclusive adoption of (N), McKenna distinguishes between *intrapersonal* theories of moral responsibility and *interpersonal* theories of moral responsibility. According to McKenna, a theory of moral responsibility is an *intrapersonal* theory just in case it does not "build an *interpersonal* dimension into an account of the very nature of responsible agency, or into the nature of responsibility for conduct" (McKenna, 2012, 45, my emphasis).

Expanding a bit, the idea is that on *intrapersonal* theories "the standpoint of responding to the agent is to be understood as entirely derivative and as concerned first and foremost with getting straight the more basic independent responsibility facts about the agent" (McKenna, 2012, 45). Alternatively, a theory of moral responsibility is an *interpersonal* theory just in case it "explains responsible agency in terms of interpersonal relationships essential to the nature of such agency [...and] it explains moral responsibility for conduct in terms of the possibility of interpersonal transactions between the responsible agent and those holding responsible" (McKenna, 2012, 45). Combining reasons in favor of an interpersonal theory of moral responsibility with a version of a

---

appropriateness of holding responsible. Thus, the first route does not provide us with a reason to adopt (N) any more than it provides us with a reason to adopt a version of (C) where no reference is made to holding responsible. The disambiguation of (N) does nothing to cure this failing. The second route helps to support *any* account that derives what it takes to be morally responsible from our best theory about holding responsible. Disambiguating (N) does nothing to secure it as *the only* viable option for building a theory of being morally responsible that is grounded in our best theory of holding responsible. Thus, the disambiguation does not strengthen the second route for (N).

reactive attitudes account of responsibility responses, McKenna comes away with (N).

McKenna offers little in the way of supporting the adoption of an interpersonal theory of moral responsibility or his reactive attitudes account. But, we can nevertheless grant him each. The trouble for McKenna is that whether we take up those theses or not, neither the truth of (N) nor RD-Normativism is supported.

Adherence to the view that only an interpersonal theory of moral responsibility and only a reactive attitudes account can be correct supports nothing more than is supported by the adoption of the position that only a moderate metaphysical theory can be correct together with a restriction on what holding responsible amounts to. As I showed above, adopting the position that only a moderate metaphysical theory can be correct supports the view that the only correct theory of moral responsibility is one that derives what it takes to be morally responsible from a view about the appropriateness conditions of holding responsible—i.e. The Derivation Thesis. And, there are viable examples of views that do this but do not adopt (N) in order to do it.

Marianna Oshana, for instance, adopts a view on which “a person is responsible for an act if and only if it ought to be the case that the person account for her behavior, where doing so involves giving some statement of the person’s beliefs or intentions regarding the act” (Oshana, 1997, 77). Thus, Oshana explicitly denies (N). But, even McKenna admits that her view is an interpersonal theory, due to the fact that for her “agency is pinned to the possibility of social transactions with members in a moral community prepared to hold responsible” (McKenna, 2012, 46).

Some, including McKenna, might object to this example, since Oshana appears to

deny the view that a reactive attitudes account must be correct.<sup>41</sup> But, it should be clear that one could adopt Oshana's bi-conditional and a reactive attitudes account, since her bi-conditional says nothing about the nature of responsibility responses. So, the objection is short lived. We have, then, a clear case of an alternative to (N) and RD-Normativism that is both interpersonal and compatible with reactive attitudes accounts.<sup>42</sup>

Furthermore, and perhaps more importantly, we can outline what it takes to stay true to the precepts of an interpersonal theory and a reactive attitudes account. First, one must determine what it takes for some act to be attributable to an individual as the basis of moral appraisal as well as what it takes for positive and negative appraisals to be appropriate.<sup>43</sup> Second, derive from the results of that investigation what it takes to possess morally responsible agency. Finally, argue for the view that the attribution and/or the appraisals that we are concerned with are conative-affective attitudes. No need for (N), and trivially, no need for citing (N) as fundamental or the only correct analysis of being morally responsible either.

We can get along, then, without citing (N) as the only correct analysis of being morally responsible or RD-Normativism and still meet the basic constraints laid out in

---

<sup>41</sup> For perspicuity, I'm not a proponent of reactive attitudes accounts either. See, my (unpublished m.s.).

<sup>42</sup> Another case would be those views developed by Scanlon (1998, Chpt. 6; 2008, Chpt. 4, esp. 128 and 198 – 204), A. M. Smith (2000; 2005; 2007; 2008; 2012), and Shoemaker (2011). Roughly, the general idea is that one is morally responsible for  $\phi$  just in case  $\phi$  is attributable to her as the basis of moral appraisal. And, according to Scanlon (2008, Chpt. 4, 128), the reasons supporting such an attribution flow from our interpersonal relationships. A. Smith is silent about the source of the reasons supporting such attributions, but could adopt Scanlon's view. Shoemaker breaks rank in that he identifies what he sees as three various types of moral responsibility. Still, S's being answerable for that for which S is being held responsible and the view that (at least some) of the reasons for holding one responsible flow from the relationships in which we are involved both feature prominently in his view. And, like Oshana's view none of the three adopt (N) as the only correct analysis of being morally responsible. Shoemaker (forthcoming (a)) thinks (N) is true for at least one of his types of responsibility—namely, accountability responsibility.

<sup>43</sup> See, e.g., Scanlon (1998, Chpt. 6; 2008, Chpt. 4) and A. Smith (2000; 2005; 2007; 2008).

the two traditional routes to that view. But, as was pointed above, normativists take themselves to be developing a better version of the picture of moral responsibility painted by Strawson. So, we might wonder whether we can develop an account of moral responsibility that is robustly Strawsonian without holding (N) as the only correct analysis of being morally responsible. In the next section, I show that the answer is affirmative.

### 3: STRAWSONIAN THEMES

In some ways it should not be surprising that we can retain a robustly Strawsonian account of moral responsibility without positing (N) as the only correct fundamental analysis of being morally responsible. Given that the reasons just canvassed that are traditionally offered in favor of (N) strongly track general themes in Strawson's original work, I have already covered ground important to showing that we do not need to exclusively adopt (N) or RD-Normativism in order to stay true to them. Furthermore, some of the themes aptly employed by Strawson can be found prior to publication of his essay.

For instance, the idea that being morally responsible should be derived in large part from appropriateness conditions of holding responsible is alive in French sociologist, Gabriel de Tarde's (1912, 46 – 49) work. Also, the connection Strawson draws between moral responsibility and sentiments is developed elsewhere by Bishop Joseph Butler (1726) and David Hume (1777, 125 – 136).<sup>44</sup> Further, the keystone idea that we oscillate

---

<sup>44</sup> Wallace (1994, 8n.) also points to Bishop Butler's work.

between seeing individuals as participants in the moral community and not seeing them as such is skillfully employed by Elizabeth Lane Beardsley (1960). And, she uses that idea to make points similar to those made by Strawson without the explicit use of the reactive attitudes. But, none of this is really enough to show that we can develop a view of moral responsibility that is as rich as Strawson's without taking an RD-Normativist approach according to which (N) is the only correct fundamental analysis of being morally responsible. In order to show this, we have to look closely at important points of interest in Strawson's work as those attracted to normativism have. To do this, I revisit these points as an indispensabilist—namely, one who agrees with the general perspective that the interpersonal nature of our practice of responding to individuals as morally responsible is indispensable to a correct explanation of being morally responsible. Furthermore, I revisit these points of interest as an indispensabilist concerned with grasping these themes not only at a normative level, but also at a non-normative level.<sup>45</sup>

We have already seen that incorporating an interpersonal account of moral responsibility and an account of responsibility responses according to which instances of blame and praise are constituted by motivationally efficacious attitudes (like conative-affective attitudes) is required in order to maintain a Strawsonian picture of moral responsibility. So, I will discuss three other main points of interest of that picture.

First, the account must retain the idea that moral responsibility centers on “the very great importance that we attach to the attitudes and intentions towards us of other human beings, and the great extent to which our personal feelings and reactions depend

---

<sup>45</sup> I do not mean for this distinction to indicate that the normative and non-normative are metaphysically distinct. For my purposes we could replace non-normative with descriptive or merely descriptive.

upon, or involve, our beliefs about these attitudes and intentions” (Strawson, 1962, 75). That is, the account must “emphasize how much we actually mind, how much it matters to us, whether the actions of other people—and particularly some other people—reflect attitudes toward us of goodwill” (Strawson, 1962, 76). In a similar vein, Beardsley writes:

We value in a special way those whose acts meet the standards of moral worth and moral credit, and this is something we cannot change. As Spinoza saw, it is true—even in a determined universe—that “we desire to form for ourselves an idea of man upon which we may look as a model of human nature.” The idea of a man who performs a right act voluntarily, knowingly, and from a good desire, and the idea of a man who, when confronted by odds, can still do these things—these *are* the models we have formed. Conformity to these patterns is what we regard as worthy of praise, and deviation from them in certain ways is what we regard as worthy of blame. We cannot feel about persons who thus conform or deviate as we do about animals or inanimate objects which measure up or fail to measure up to certain other standards. All this being so, judgments of praise and blame based on moral worth and moral credit are not only legitimate but vitally necessary parts of moral discourse. They are answers to questions that we cannot help asking. (Beardsley, 1960, 14, her emphasis)<sup>46</sup>

In general, this theme can be referred to as *the quality of will theme*.

Broadly, the quality of will theme is simply that the attitudes involved in responsibility responses are tracking the quality of the will underwriting an individual’s attitude or action.<sup>47</sup> Seemingly, there are numerous ways of incorporating this theme into a theory of moral responsibility.<sup>48</sup> Rather than list various ways one might incorporate

---

<sup>46</sup> The quote from Spinoza is cited by Beardsley (1960, 14 n. 18) as: B. Spinoza, *Ethics* (Oxford Press, 4<sup>th</sup> e.d, 1930), p. 179.

<sup>47</sup> I do not mean for this to rule out a pluralistic view about what ‘quality of will’ comes to. For details on such a view, see Shoemaker (forthcoming (a) and (b)). I employ Shoemaker’s pluralism in Chapter 5.

<sup>48</sup> See Shoemaker (forthcoming (a) and (b)).

the quality of will theme, it is enough for my purposes to show that it is nothing about (N) itself nor positing (N) as fundamental that guarantees that a theory will incorporate this theme.

Notice, for example, that capturing the quality of will theme involves adopting a particular view about the nature of responsibility responses. Clearly, since (N) simply tells us what it is to be morally responsible and is silent about the nature of responsibility responses, one can hold or deny (N) without taking on any particular view about the nature of such responses. Of course, it hints that the nature of holding responsible is such that instances of it can be evaluated for appropriateness. But, that is not nearly enough to commit one to a theory of responding to an individual as morally responsible that stays true to the quality of will theme. Further, adopting the view that (N) is the most fundamental analysis of moral responsibility suggests that explaining being morally responsible requires only role-invoking explanations. But, we can find a place for non-role-invoking explanations as well in the quality of will theme.

The quality of will theme gives us at least one part of the appropriateness conditions on holding responsible. Very roughly, the extent to which the attitudes involved are appropriate is the extent to which they track manifestations of will and qualities of them and motivate in light of both. In other words, whatever the final account of the appropriateness conditions on responsibility responses, those conditions will hang tightly with the notion that the attitudes involved are sensitive to the qualities of will evinced by oneself or others.<sup>49</sup> Thus, when asked what it is to be morally responsible

---

<sup>49</sup> See, e.g., Scanlon (1981; 2008, Chpt. 4, esp. 127 – 128), Wallace (1994, Chpt. 5, esp. 127 – 136), Arpaly & Timothy Schroeder (1999, esp. 175), Arpaly (2003, Chpt. 3, esp. 79 – 93, and Chpt. 5; 2006, Chpt. 1, esp. 14 – 17), Heironymi (2004), Shoemaker (2011, esp. 619 – 624), and McKenna (2012, Chpt. 3).

or worthy of moral-responsibility responses, aligned with the quality of will theme we could explain that it is a function of being capable of manifesting a will with a particular quality and/or exercising that capacity. Here, then, we are giving an explanation that elucidates the grounds for worthiness of a certain sort without invoking the role such worthiness plays because we were asked for an explanation of the source of normativity. Moreover, the explanation is clearly compatible with (N).

A second important theme is related to the quality of will theme. Strawson draws a tight connection between instances of responding to an individual as morally responsible and our expectation or demand that individuals bare good will toward ourselves and others, especially those with whom we stand in special, deeply personal relationships.<sup>50</sup> The connection, which is spelled out quite well by Wallace, is that our disposition to respond to the behavior of others is grounded by expectations and/or demands that our wills carry a certain quality.<sup>51</sup> Call this *the demand theme*. Once again, it should be clear that it is nothing about (N) itself or holding that (N) is the only correct analysis of being morally responsible that guarantees that a theory will incorporate the demand theme, since (N) says nothing about what grounds our disposition to hold responsible. The demand theme is really just about the psychological underpinnings of our practice of responding to ourselves and others morally responsible.<sup>52</sup> Thus, whether we take up (N), (N) as the only correct fundamental analysis of being morally responsible, or RD-Normativism is not to the point at all.

---

<sup>50</sup> Strawson (1962, esp. 76 – 77 & 84 – 85). See, also, Beardsley (1960, 4 – 7 & 14 – 17).

<sup>51</sup> Wallace (1994, esp. Chpt. 2 and Chpt. 3).

<sup>52</sup> For an interesting development of this idea along Strawsonian lines, see Shaun Nichols (2007).

The final theme I look at is connected to both the quality of will theme and the demand theme and can be usefully divided into two sub-themes. The general theme, to be spelled out presently in more specific terms, is that instances of holding responsible are connected in interesting and illuminating ways. Call this *the connection theme*. One connection (and sub-theme) cited by Strawson is what he identifies as *the logical connection*, and the other connection he calls *the human connection*.<sup>53</sup>

Consider first the logical connection. Of this theme, Strawson writes:

The personal reactive attitudes rest on, and reflect, an expectation of and demand for, the manifestation of a certain degree of goodwill or regard on the part of other human beings towards ourselves; or at least on the expectation of, and demand for, an absence of the manifestation of active ill will or indifferent disregard. [...] The generalized or vicarious analogs of the personal reactive attitudes rest on, and reflect exactly the same expectation or demand in a generalized form; they rest on, or reflect, that is, the demand for the manifestation of a reasonable degree of goodwill or regard, on the part of others, not simply towards oneself, but towards all those on whose behalf moral indignation may be felt, i.e., as we now think, towards all men. The generalized and non-generalized forms of demand, and the vicarious and personal reactive attitudes which rest upon, and reflect, them are connected [...] logically. (Strawson, 1962, 84)<sup>54</sup>

The logical connection, then, is simply the connection spelled out in the demand theme. And, so, delimits what Strawson sees as the psychological underpinnings of responsibility responses. Thus, one could accept or reject (N) without disagreeing with the logical connection sub-theme of the connection theme.

---

<sup>53</sup> More accurately, Strawson first cites a phenomenon and draws the conclusion that the reactive attitudes and demands are connected logically. Second, he claims that reactive attitudes bear a human connection.

<sup>54</sup> For Strawson, the personal reactive attitudes he refers to are attitudes like guilt, resentment, self-approbation, and gratitude. Their 'vicarious analogs' are indignation and approbation.

We find similar results regarding the human connection sub-theme. Concerning the human connection, Strawson writes:

All these three types of attitude [i.e. the personal reactive attitudes and their vicarious analogs] are humanly connected. One who manifested the personal reactive attitudes in a high degree but showed no inclination at all to their vicarious analogues would appear as an abnormal case of moral egocentricity, as a kind of moral solipsist. Let him be supposed fully to acknowledge the claims to regard that others had on him, to be susceptible to the whole range of self-reactive attitudes. He would then see himself as unique both as one (the one) who had a general claim on human regard and as one (the one) on whom human beings in general had such a claim. This would be a kind of moral solipsism. But it is barely more than a conceptual possibility: if it is that. In general, though within varying limits, we demand of others for others, as well as of ourselves for others, something of the regard which we demand of others for ourselves. Can we imagine, besides that of the moral solipsist, any other case of one or two of these three types of attitude being fully developed, but quite unaccompanied by humanity—a moral idiot or a saint. For all these types of attitude alike have common roots in our human nature and our membership of human communities. (Strawson, 1962, 85)

As we can see, the human connection sub-theme like the logical connection is simply about our psychology. Here again, then, we find the participant stance, but we also find our nature as social creatures featured prominently.<sup>55</sup> According to Strawson, our sympathy, empathy, or fellow-feeling explains why we typically feel and demand for others what we feel and demand of and for ourselves. And, just as with the other themes, nothing about (N) itself, its presumed fundamentality, or whether being worthy of moral-responsibility responses is a response-dependent property ensures that a theory featuring (N) will incorporate the connection theme generally or its sub-themes more particularly.

---

<sup>55</sup> I owe much of my thought here to John Diegh (2011).

Perhaps there are other strongly relevant themes coursing through the Strawson's sketch of moral responsibility. Still, we should now see that neither a commitment to (N), the exclusivity of (N) as *the* analysis of being morally responsible, nor RD-Normativism is itself entailed by a commitment to any of the more substantive views strongly suggested by that sketch. And, none of those Strawsonian themes entail any of those normativists views.

#### 4: CONCLUSION

Normativists lean on the indispensability-thought in order to support their view and some lean on it to support the view that (N) is the only correct fundamental analysis of being morally responsible. But, nothing the members of this latter sub-set of normativists say gives us reason to believe that their view is the best implementation of that thought nor that we even need to adopt (N) in order to stay true to it. Perhaps this is the basic line of thought underlying skepticism of normativism. Oshana conveys the skepticism nicely:

How we naturally act toward persons we believe are responsible offers evidence of just that—behavior to which we are arguably predisposed and nothing more. But why should the fact (if it is a fact) that we are committed to these attitudes justify a belief that people are responsible for their acts? We do not call a person responsible simply because she meets certain conditions for free moral agency, nor because others regard her as if she were a free agent anymore than we call an emerald priceless because we value or desire it. Rather, we call a person an appropriate subject of the reactive attitudes because the person is responsible. And that judgment [...] is explained by the fact that the agent is accountable. (Oshana, 1997, 75)

Nomy Arpaly has similar reservations:

The moral emotions are indicators of concern for morality, which is in turn the attitude of the person who blames or approves. But, contrary to those who think that blameworthiness and praiseworthiness are epiphenomena of the reactive attitudes, it is the blameworthiness and praiseworthiness of the agent for her action that makes her an “appropriate” object of these emotions, and not the other way around. (Arpaly, 2006, 28)

I largely agree with the thoughts expressed by skeptics like Oshana and Arpaly, but their skepticism is too sweeping. What they overlook is that we can be normativists and think that there are analyses of being morally responsible that only cite the conditions under which responsibility responses would be fitting or appropriate. Thus, aligned with normativists and individuals like Oshana and Arpaly we can accept that we are accidents waiting to happen for whom the quality of an individual’s will is a (pro tanto) reason for us to respond in specific ways thereby rendering the individual morally responsible (for certain attitudes and behaviors). And, when asked, depending on the circumstances, we can explain this stance by employing a role-invoking or non-role-invoking explanation. Furthermore, it seems each will be suitable and no less true or fundamental than the other.

From here, then, as a normativist I turn my focus to accounts of responding to an individual as responsible. In the next chapter, I offer an account of responsibility responses that matches their shape better than the morally important sentiments Strawson identified as reactive attitudes. Thus, I continue to sketch my unorthodox Strawsonian picture that I will employ as a strong foundation for rejecting global skepticism about moral responsibility.



### III: RECOGNITION AND RESENTMENT

Responding to each other as morally responsible is central to our practical lives. Judging that a friend has wronged his partner pushes the judge to reinterpret her friend and perhaps even primes her to reproach him. Often, though not always, these sorts of judgments are accompanied by affect and often that affect is felt. Hence the popularity of the view attributed to P. F. Strawson—namely, the

*Reactive Attitudes Account:* Responding to an individual as morally responsible is constituted by a complex conative-affective attitude (i.e. a reactive attitude like resentment, indignation, guilt, gratitude, approbation, and pride).<sup>56</sup>

Recently, pressure has been put on proponents of this view to explain cases where no affect is present. The thought is that it would be nice to capture in a unified account the practical and emotional import that affective and non-affective responsibility responses like praise and blame have. This paper attempts to provide that.

The key, I take it, is to reject the Humean idea that cognitive attitudes cannot motivate absent aid from independent desire. On my anti-Humean reconstruction of the Reactive Attitudes Account, an individual's recognition that someone has manifested a will of a particular quality is enough to spur the change in (or reinforcement of) the recognizer's behavioral, attitudinal, and interpretive dispositions characteristic of praise and blame. For ease of prose, I refer to the account as the *Recognition Account*.

The paper unfolds in three parts. First, I delineate an account of blame explicitly built around the Humean idea that cognitive attitudes are essentially inert without help

---

<sup>56</sup> Strawson (1962). Restricting the set of reactive attitudes will not effect my arguments here.

from an independent desire. We should, I argue, reject the Humean Account, since it distorts the phenomenon and overlooks the resources available to Strawsonians (§1). Second, I follow up by deconstructing the reactive attitudes and rebuilding the Strawsonian view through an anti-Humean lens (§2). I close the section with a brief defense of the anti-Humean moral psychology central to the Recognition Account. Finally, I further develop the picture by attending to the concern that views like mine ‘take the blame out of blame’ (§3). This development reveals that the Humean Account returns strange results when considering how it must construe the connection between blame and desert. We also see that the anti-Humean alternative smoothly accommodates a robust sense of blame as ‘something the wrongdoer has coming to them’. Indeed, a more wholeheartedly anti-Humean embrace of the Strawsonian idea that responding to each other as morally responsible is, at bottom, simply about motivationally efficacious recognition illuminates the contours of our responsibility responses better than extant accounts of the phenomenon. Hence, we see that the Recognition Account is superior to the Strawsonian model while nevertheless preserving its beauty. I turn, then, to begin my push against the idea that responsibility responses must be aided by independent attitudes beyond the evaluative judgments central to them.

## 1: BEYOND BELIEF?

Many are dissatisfied with the once popular

*Ledger Account:* A responsibility response is constituted by a belief that an individual, S, was morally responsible for  $\phi$  and a

positive/negative mark on S's ledger of life.<sup>57</sup>

As George Sher (2006, 75 – 76) notes, the metaphysical commitments regarding, e.g., the existence of *ledgers of life* and the positive/negative marks on them, might be mitigated by construing the Ledger Account as merely positing beliefs or judgments with particular contents. Making this move imparts upon the ledger account the virtue of being focused on a “complex state of mind or attitude rather than an act” (Brandt, 1958, 8). In light of this psychological focus, it is widely taken to do better than the

*Expression Account*: A responsibility response is constituted by the expression of an attitude by the responder toward the respeegee.<sup>58</sup>

Though the Ledger Account does better than the Expression Account insofar as it attempts to illustrate the psychological phenomena involved, it is rejected by many. The primary grounds for rejection are that the Ledger Account does not do justice to the thought that responding to individuals as responsible for their behavior plays a central role in our *emotional* lives.<sup>59</sup> Thus, even if the Ledger Account is made more metaphysically acceptable, it faces the concern that it fails to make sense of *emotive-*

---

<sup>57</sup> Supporters are, e.g., Feinberg (1970, 124 – 125), Glover (1970, 64), Zimmerman (1988, 38), and Haji (1998, 8).

<sup>58</sup> See, e.g., Smart (1961). Smart's view is often mistakenly taken to be bound to utilitarianism. This is likely due to Smart's endorsement of consequentialism, which leads him to tie the appropriateness of blame to consequences (Arneson (2003)). But, as Beardsley (1969, 34) explains, the Expression Account does not entail utilitarianism, and, as Arneson points out, utilitarianism does not entail it either. Still the mistake continues. For instance, Saul Smilansky conflates the account with a compatibilist story and attributes it to Hobbes, Mill, Sidgwick, Schlick, Stevenson, Nowell-Smith, and “the early Ayer, etc.” (Smilansky, 2000, 27). For examples of the wide consensus against the Expression Account, see Beardsley (1957; 1978, 573), Glover (1970, 56 – 57), Zimmerman (1988, 38), Wallace (1994, 56 – 61), Hieronymi (2004), and Sher (2006, 74).

<sup>59</sup> See, e.g., Wallace (1994, 52 – 62), (Sher, 2006, 77), and McKenna (2012, Chpt. 2 esp. 45 – 46).

*centricity*. An acceptable account of responsibility responses, then, ostensibly must go beyond cognitive attitudes like beliefs to capture the complex psychology involved.

A key motivation for the Reactive Attitudes Account is that it makes good sense of emotive-centricity, or the central role that holding responsible plays in our emotional lives. I argue, following Sher, however, that it is a mistake to secure emotive-centricity by attaching it to a particular type of attitude that is *by its nature* affective. Interestingly, the Humean account Sher offers as a replacement suffers from a similar fault insofar as emotive-centricity arrives on the scene by way of certain frustrated desires. These problems are illuminated by considering certain cases. The cases reveal that purchasing emotive-centricity by embracing an account on which emotion or desire frustration is (partly) constitutive of responding to individuals as fit for normative appraisal costs proponents a unified and accurate story of the phenomenon.

### 1.1: Blame without Affect

Consider Sher's cases of affectless blame:

We may, for example, feel no hostility toward the loved one whom we blame for failing to tell a sensitive acquaintance a hard truth, the criminal whom we blame for a burglary we read about in the newspaper, or the historical figure whom we blame for the misdeeds he performed long ago. As the latter examples suggest, blaming is something that we can do regretfully or dispassionately and that need not be accompanied by any rancor or withdrawal of good will. At least offhand, it seems perfectly consistent to suppose, first, that the stance I take toward my daughter for shading the truth about how much of her homework she has done is genuine blame, but, second, that that stance involves no modification at all “of the

general demand that another should, if possible, be spared suffering". (Sher, 2006, 88)<sup>60</sup>

Whether any of these responses are fitting is beside the point. Each appears to be an instance of responding to an individual's behavior as if she were normatively responsible for it where those responding are doing something more than simply noting that objectionable behavior occurred. Sher's idea is that those who blame are ensconced in their recognition acts of cowardice, injustice, or disrespect in such a way that certain of their behavioral or attitudinal dispositions shift. This sort of recognition-change marks the responses as blame. And, when carrying no affect the combination is a case of *affectless* blame. So, Sher can recognize disrespect in what his daughter has done and be moved to interpret her or interact with her in a manner resonate with that recognition, and yet feel no anger, hostility, or resentment toward her. Since proponents of the Reactive Attitudes Account are committed to the view that a response to a manifestation of will involving anything less than an essentially affective reactive attitude is *not* a responsibility response proper, they seem to face a dilemma. They must either deny that such episodes are instances of blame or sacrifice uniformity to explain the cases.

Some reactive attitudes theorists do dig in and hold that there *must* be some felt affect involved in order for a mental episode to be a responsibility response, even if the affect is so muted that feeling it requires some sort of acute perception.<sup>61</sup> Others claim that cases like Sher's involve a *different way* of responding to an agent's act as one for which she is responsible and negatively appraisable. On their view, the responder has a subjunctive belief, e.g., the belief that it *would* be appropriate for her to have a reactive

---

<sup>60</sup> Sher is quoting Strawson (1962, 90). Cf. Wallace (1994, 76; 2011) and McKenna (2012, 24 – 26).

<sup>61</sup> See, e.g. Wertheimer's (1998).

attitude *were* she to have such an attitude.<sup>62</sup>

But, these ways of understanding cases of affectless-blame basically just deny what seems clearly true, *viz.* to respond to an individual as if she were responsible and worthy of negative appraisal for what she did without feeling any affect is to blame them essentially in the same way that resenting someone is. Each recognizes that someone behaved in a particular manner. And, each involves corresponding changes in the way the responding individual is disposed to understand and interact with the world. I see no reason to treat the affectless cases of blame as anything but genuine and central cases of the phenomenon that the Reactive Attitudes Account is supposed to capture. If the appeal to the reactive attitudes fails to explain those cases, that is a significant mark against it. However, as my deconstruction of reactive attitudes and anti-Humean rebuild will show we can preserve the Strawsonian account of how responsibility responses involve more than description. First, I turn to discuss Sher's Humean alternative.

## 1.2: Blame Without Desire

Proponents of the Humean alternative face a dilemma similar to the one seemingly facing proponents of the Reactive Attitudes Account. Recall that the latter are ostensibly met with the choice between (a) denying that a particular type of response counts as a responsibility response and (b) adopting a non-uniform account of such responses. The Humean alternative arises as an attempt to find a place for the emotive-centricity of those responses without appealing to certain affective attitudes. Roughly, the Humean locates

---

<sup>62</sup> See, e.g. Wallace (1994, 75 – 78) and McKenna (2012, 33 & 106). McKenna considers adopting the position that holding responsible requires some felt affect muted or otherwise.

emotive-centricity in a particular kind of desire. Once we see that there are responsibility responses in the absence of such desires, we see that the Humean's options are the same as the Strawsonian's: either explain those cases at the cost of denying what seems true or explain them at the cost of unification.

Consider Sher's Humean view. Sher supplants the Reactive Attitudes Account with

*Humean Blame:* A responsibility response like blame is constituted by a belief that someone, S, is blameworthy for  $\phi$ -ing and the frustration of either (a) a desire that S not have been that way (e.g. the way S was that made  $\phi$ -ing seem like a reasonable thing to do) or (b) a desire that S not have  $\phi$ -ed.<sup>63</sup>

On Sher's view, the desires cited in (a) and (b) are aligned with the moral system to which the one desiring subscribes. For Sher, those desires explain the connection between blame and certain behavioral and attitudinal dispositions characteristic of it. It is the frustration of those desires that motivate these dispositions.<sup>64</sup> So, for instance, suppose Nell's desire that Jaclyn not have wrongfully manipulated Dario into a risky home loan is frustrated and so Nell blames her for it. On Sher's picture, this leads Nell to take up Jaclyn's perspective and see that Jaclyn did not take the fact that her action would treat Dario disrespectfully as a reason to refrain from doing it. This realization, then, disposes Nell to rebuke or feel indignation toward Jaclyn.

Sher attempts to make sense of affectless-blame as an instance of blame where

---

<sup>63</sup> To be clear, this is not to be confused with the view often attributed to Hume according to which we blame individuals for their characters but not their actions. However, Sher (2006, Chpts. 2 – 4) offers an excellent discussion of that view and incorporates a version of it into his picture of blame. [omitted]

<sup>64</sup> Sher (2006, 93 – 95 & 119 – 133).

one of two things occurs. Either the dispositions to have affective attitudes simply fail to manifest or the only dispositions that manifest are non-affective. For Sher, the failure to manifest affective attitudes is due to mitigating factors like the fact that the individual's bad behavior occurred in the very distant past or the fact that this is his daughter's first time being dishonest about schoolwork. The idea, then, is that responsibility responses play an important role in our emotional lives but external factors may silence their affective attributes. Thus, for Sher, emotive-centricity attaches to negative responsibility responses like blame by way of desires with a particular content and their frustration.<sup>65</sup>

My concern is that Humean Blame overcorrects in attempting to account for affectless blame and is thereby committed to a less plausible story about affectful blame. Recall that on Humean Blame all blame just is a combination of a certain belief, a certain desire, and the frustration of that desire. Here, that an individual does not always feel when she blames might make sense due to the fact that the belief and desire components are not essentially affective. In fact, unmanifested, none of the blame-characteristic dispositions that are presumably imbued upon the individual bearing the requisite blame-constituting belief-(frustrated)-desire combination are essentially affective either. But, affectful blame seems to first and foremost involve emotions which simply are not

---

<sup>65</sup> Generalized, I think the account would look something like this:

A responsibility response is constituted by a belief about an individual, S, (e.g. that some agent, S, is blameworthy for  $\phi$ -ing) and either (I) the frustration of a desire (e.g. that S not be as they were, that S not have  $\phi$ -ed, etc.) or (II) the satisfaction of a desire (e.g. that S be as they were, that S had  $\phi$ -ed).

Really, to adequately sharpen the view, we need to know where it stands in the family of views related by the fact that each accepts some version of the Humean picture of moral psychology. For an excellent codification of that picture, see Barry (2005, 197). My argument here doesn't hinge on these fine points. For an argument for the Recognition Account along those lines, see [omitted].

constituted by a belief-(frustrated)-desire combination.<sup>66</sup> We may, for instance, feel searing hostility toward our loved one, David, whom we blame for failing to tell our sensitive acquaintance, Maria, the hard truth that it was his friends who accidentally poisoned her not the person currently charged with the crime. We may likewise feel intense enmity toward the officer who we see choke an unarmed man despite his dying words “I can’t breathe”.

Furthermore, these passionate responses can be quite complex. Imagine, for instance, that after being promoted by phone, Amber met with the board for lunch. The next day, her boss, Mr. Benson, calls her up to his office. Benson explains that he spoke with the board and agreed to cancel the contract. “I know we had a deal, and I’m afraid I must renege. The board is concerned. They think that the, um, the face of the marketing department should be, um, uh, well more like Don Driver’s.” As Benson goes on trying not to say what is clearly being said, Amber hears the rationalizations harkening the specter that’s always haunted her—the phantom of being a black, capable woman in America. Benson continues, but, Amber’s resentment is fueling a fire that’s turning to white hot indignation, washing the company in a repulsive hue that mutes or erases its once brilliant shine of opportunity. Previously disposed to smile and shake hands, Amber feels like grabbing Benson by the collar to bring his face to hers, make him say what he was trying so hard not to say, *make him announce it*. His promise that she is first in line for ...crumbles as it creeps to her ears. Nothing could mitigate the fact that for Benson and the company one’s race is grounds for divorce. Only her battle-hardened ability to remain prudent in the face of injustice keeps her seething from reaching the surface as

---

<sup>66</sup> For starters, see, e.g., Lewis *et al* (2010, Chpt. 1).

she shakes hands and walks to the sanctuary of the elevator.

As these examples suggest, sometimes we blame without regret, full of passion and rancor. I have no doubt that Sher can spin a Humean story to explain each situation. Certainly, e.g., Amber has the cognitive component—recognition of ill will and/or a belief that a wrong has occurred. And, perhaps in the background Amber has a desire that those who wronged her wouldn't have done so. We might even imagine that Amber took up the perspective of those who frustrated her desire, realized that they did not see that one's race is not acceptable as a reason for breaking contracts, and thereby was inundated with blame-characteristic dispositions which manifested in unexpressed emotions that reorganized her perception and interpretation of aspects of her world.

That the Humean has a tale to tell about affectful blame notwithstanding, it never tells the most obvious tale—namely, the story where the recognition is embodied in a conative-affective reactive attitude or emotion completely constitutes the episode of blame. However, the Reactive Attitudes Account can. Moreover, so can the Recognition Account while also doing what the Reactive Attitudes Account cannot—namely, make sense of cases of affectless blame without denying their reality or giving up on a unified view of responsibility responses. The only cost is rejecting the Humean idea that cognitive attitudes cannot motivate absent aid from independent desire. I'll discuss what that cost really comes to after sketching the account I have in mind.

## 2: THE RECOGNITION ACCOUNT

The Recognition Account aims to preserve the beauty of the Reactive Attitudes Account

while avoiding the errors that beset that view and Humean Blame. A guiding thought is that emotive-centricity should not be located in essentially affective attitudes nor the frustration of purely conative attitudes. But, rather than follow the Humean and search beyond belief and beyond the reactive attitudes I deconstruct the latter and identify two causal-functional roles traditionally set aside for them. I call these roles, *the recognition function* and *the emotion function*. I follow by building the account around them.<sup>67</sup> The result is an anti-Humean view friendly to the general Strawsonian picture. After sketching the account, I attend to arguments against anti-Humean attitudes. That discussion will further illuminate the anti-Humean shape of responding to individuals as worthy of normative appraisal for their behavior and it will serve to show that the cost born by taking up the Recognition Account is not as high as some might think.

## 2.1: Recognition

Strawson introduces his version of the Reactive Attitudes Account in a famous attempt to reconcile two views—namely, *pessimism* and *optimism*. Here, the *pessimist* thinks that the coherence and justification of our moral concepts and practices would be lost in the face of determinism while the *optimist* holds that they live on albeit colored by consequentialist thinking.<sup>68</sup> Key pieces of Strawson’s reconciliation are the “non-detached attitudes and reactions of people directly involved in transactions with each other; of the attitudes and reactions of offended parties and beneficiaries; of such things

---

<sup>67</sup> McGeer’s (2013) account of blame is similar insofar as it specifies blame by reference to causal-functional roles. However, on McGeer’s account blame remains an emotion. [omitted]

<sup>68</sup> Strawson (1962, 72 – 75). Cf. Beardsley (1960).

as gratitude, resentment, forgiveness, love, and hurt feelings” (Strawson, 1962, 75)—i.e. *the reactive attitudes*. Roughly, Strawson’s argument is that the reactive attitudes form the core of our moral responsibility practices and their propriety does not turn on the truth or falsity of determinism.<sup>69</sup> Rather, reactive attitudes are sensitive to the quality of will manifested by an individual’s actions and, so, are *only* inappropriate when directed toward one who is excused or exempt.<sup>70</sup> Additionally, for Strawson, roughly, only the behavior of those taken to have a relevant excuse or taken to be outside the moral community are unfit subjects of the reactive attitudes insofar as their behavior does not (properly) manifest a moral attitude, good or ill. Finally, since Strawson holds that being causally determined does not number among the relevant excuses and has nothing to do with being taken to be outside the moral community, our moral responsibility practices survive.<sup>71</sup>

Whether Strawson’s argument is sound is beside the point. The point is that the reactive attitudes are attitudinal modifications of a certain sort. In general, these attitudinal modifications track the quality of will manifested by the decisions reflected in one’s behavior.<sup>72</sup> The reactive attitudes are paradigm cases of attitudes that are sensitive to such things. Even so, *cognitive* attitudes like evaluative beliefs or judgments can serve this function as well. Call it, *the recognition function*.

---

<sup>69</sup> Strawson (1962, 77 – 80).

<sup>70</sup> Strawson (1962, 80 – 83). ‘Quality of will’ is notoriously difficult to unpack. I tend to think of it in terms of the excellence or non-excellence of one’s decision where that quality is set by various aspects of the agent’s varying moral attitudes. Cf. Paul C. Taylor (2004, 33 – 35), Arpaly (2006, 14 – 15), and McKenna (2012, 58). I am very sympathetic to Shoemaker’s (forthcoming a, b) pluralistic.

<sup>71</sup> Strawson (1962, esp. 77 – 83 & 87 – 89).

<sup>72</sup> See, e.g., Strawson (1962), Scanlon (1988; 1998, 276 – 294; 2008, esp. 127 – 138), Wallace (1994, esp. 127 – 136), Hieronymi (2004), Helm (2011), and McKenna (2012, Chpt. 3). Cf. Arpaly (2003, Chpt. 3).

On an extremely barebones characterization of the Recognition Account, then, responding to another as if she is normatively responsible for her behavior involves an attitudinal modification in the responder *in recognition of* the quality of will manifested by that individual's act.<sup>73</sup> The resulting modification can be a cognitive attitude like an evaluative belief or judgment. Thus, the Recognition Account acknowledges that a judgment like Sher's attitude toward his daughter is an instance of blame without felt affect. Here Sher simply recognizes a particular quality of will in his daughter's action but feels nothing in light of it.

Sketched thusly, the Recognition Account should be attractive to proponents of the Reactive Attitudes Account, since the attitudes involved in responsibility responses still play a key role that the Strawsonian framework assigned them to play. However, reluctance is clearly warranted, since key aspects are missing—namely, emotive-centricity and the tight connection between the recognition of will and the dispositions characteristic of such recognition. I continue, then, to deconstruct the reactive attitudes to correct for this lack.

## 2.2: Emotive-Centricity

Cases of affectless blame and affectful blame reveal that, together with the recognition of a certain quality of will, changes occur regarding the way the blamer understands and experiences her world. Recall that in blaming his daughter for her dishonesty, Sher does more than *document that* his daughter shaded the truth about her homework. And, in

---

<sup>73</sup> When holding oneself responsible, the beholder and the one beheld are simply identical.

blaming her boss and others, Amber does more than merely *note that* they have deeply disrespected her. I want to locate emotive-centricity in this elusive extra that renders the recognition involved in responsibility responses more than *mere* documentation, notation, or citation. Of course, this is what Sher and reactive attitudes theorists take themselves to be doing as well. But, I want to do it in a manner friendly to the Reactive Attitudes Account without attaching emotive-centricity to a particular type of conative or affective attitude.

To do this, I focus on Elise Hurley & Coleen Macnamara's (2010) insights regarding the Reactive Attitudes Account, since they nicely capture the idea that the recognition of will in responsibility responses goes beyond the indicative or declarative. Indeed, for them responding to an individual as if she were responsible involves "*experiencing* [that person] as having done something good, bad, right, or wrong" (*Ibid*, 2010, 386, their emphasis) rather than merely *describing* that they have. To fill this out Hurley & Macnamara adhere to the idea that responsibility responses "are forms of moral appraisal that go beyond mere beliefs" (*Ibid*, 2010, 374) that motivate a Reactive Attitudes Account by concentrating on the characteristics that such attitudes possess *as emotions*.

On their view, "to have an emotion about some feature of the world is for one's comportment in the world to change because of that feature, where that change of comportment is comprised of modifications in one's patterns of salience, tendencies of interpretation, and motivational profile" (*Ibid*, 2010, 383). Hurley & Macnamara hold that it is this emotional aspect of the reactive attitudes that explains the way in which responding to individuals as responsible is emotive-centric and so involves more than

mere description. Though I think we should resist taking up their Reactive Attitudes Account, reflecting on Hurley & Macnamara's view helps to mark out another general function that is central to responsibility responses but is typically set aside for the reactive attitudes to fulfill.

Simply put, recognizing good, neutral, or ill will is apt to motivate a change in or reinforcement of one's comportment in the world as Hurley & Macnamara describe it. So, responding to an individual as responsible for their behavior involves the causal-function of motivating a change in or reinforcement of attitudes like those that set up what is salient to the respondent, her interpretational scheme, her emotional dispositions, as well as her disposition to rebuke or praise the individual to whom she is responding. Call this *the emotion function*. And, just as belief-like attitudes can serve the recognition function, dropping the Humean idea that cognitive attitudes cannot motivate absent aid from independent desire allows that belief-like attitudes can also serve the emotion function.

Therefore, given that aptness to fulfill the emotion function partly specifies the state of holding responsible, *that state* possesses emotive-centricity not any specific type of attitude. And, rejecting the Humean idea allows that the recognition of one's quality of will, which is partly constitutive of responsibility responses, can play the other constitutive function. Thus, emotive-centricity is not lost where affect or desire (or its frustration) is absent. So, affectless-blame is seamlessly explained by the fact that one "can be in a certain functionally specified state and yet not be disposed in all of the ways that the theory says the state is apt to dispose one" (van Roojen, 1995, 47). And,

affectfull-blame is likewise explained by the fact that the recognition and emotion function are being fulfilled where resentment, indignation, or righteous anger is manifest.

So, on the Recognition Account, responding to an individual as normatively responsible for what she has done involves recognizing in her behavior a manifestation of will bearing some quality. Further, that recognition motivates a change in or reinforcement of the respondent's behavioral and attitudinal dispositional profile. Since this encompasses the roles traditionally set aside for the reactive attitudes to fulfill, Strawsonians should find the view quite appealing. Of course, one key source of resistance is that the account embraces an anti-Humean picture of moral psychology. To further support and elaborate the recognition account, then, I turn to say a bit more in defense of dropping the Humean idea that cognitive attitudes cannot motivate absent aid from independent desire.

### 2.3: Anti-Humeanism About Moral Motivation

A rejection of the popular Humean idea requires support. Though it is beyond the scope of this paper to completely defend the view, I consider and respond to the three main arguments pressed against the Anti-Humean. Since these objections make up the central cadre of reasons in favor the Humean Picture, showing that they are unsound gives us good reason to think that the anti-Humean Recognition Account is on ground no less solid than the Reactive Attitudes Account or Humean Blame.

As a preview, in a nutshell the classic worries for anti-Humean internalism about moral judgment motivation are as follows: (a) Anti-Humeanism requires a state to have

multiple ‘directions of fit’,<sup>74</sup> (b) There are cases in which belief-like states survive the absence of the desire to do as the belief-like state prescribes,<sup>75</sup> and (c) Anti-Humeanism is simply *ad hoc*.<sup>76</sup> Each concern is meant to show that the states posited by anti-Humeans are either incoherent or unmotivated. By way of diagnosing the incoherency, the conclusion that anti-Humean states are incoherent is meant to support either the view that (1) “belief and desire are modally distinct mental states with different natures” (Barry, 2004, 197) or the view that (2) in order to motivate, a cognitive attitude must combine with an independent (i.e. pre-existing, noncognitive) desire.<sup>77</sup> I answer (a) – (c) on behalf and in favor of the anti-Humean picture of moral psychology.

First, in support of (a) as a problem, Humeans hold that attitudes only have one ‘direction of fit’.<sup>78</sup> More precisely, the concern is that cognitive attitudes have a *world-to-mind* direction of fit whereas non-cognitive attitudes have a *mind-to-world* direction of fit. So, roughly, belief-like attitudes respond to the world by attempting to match the mind to it whereas desire-like attitudes push their bearer to get the world to match them. However, the objection continues, the attitudes posited by anti-Humeans both *respond to* and *push* those who possess them to make the world a certain way. The worry, then, is that such attitudes simply seem incoherent. Michael Smith writes:

Though it may sound like a coherent possibility that there should be such a state, it really isn’t [...] A state with both directions of fit would [...] have to be such that *both*, in the presence of [a perception that *not p*] it tends to go out of

---

<sup>74</sup> M. Smith (1987; 1994, 111 – 119). M. Smith attributes the basic idea of directions of fit to Anscombe (1957). However, Alvarez (2012, 66 – 71) makes a strong case against that view. Perhaps it is worth saying that the account that bears his name may not reflect David Hume’s (1777) actual view.

<sup>75</sup> M. Smith (1987; 1988, 594; 1994, 119 – 125).

<sup>76</sup> Shafer-Landau (2003, 173 n. 11)

<sup>77</sup> Again, I owe much of the phrasing here to Barry (2004, 197). And, see Barry for

<sup>78</sup> M. Smith (1987; 1994, 111 – 119).

existence, *and*, in the presence of such a perception it tends to endure, disposing the subject to bring it about that *p*. Taken quite literally, then, the idea that there may be a state with both directions of fit is just plain incoherent. (1987, 56, his emphasis)

M. Smith concludes that “what Humeans must deny and do deny is simply that agents who are in belief-like states and desire-like states are ever in a single, unitary, kind of state” (M. Smith, 1994, 119). Humean’s take the upshot to be that thesis (1) of the Humean picture—i.e., that belief and desire are modally distinct mental states with different natures—is accurate.

Though widely wielded, M. Smith’s argument does not establish what it sets out to establish. Since the belief-like portion of the mental state championed by anti-Humeans has a different content than the desire-like portion, it is not surprising that one state might persist where the other does not. As Huw Price (1989) makes clear, this sort of argument

...overlooks the possibility that if there were a propositional attitude with alternative representations as a belief and as a desire, its content might be different in each case. Indeed its content would have to be different, for no one is suggesting that for a fixed *p* BEL(*p*) [(i.e. a belief that *p*)] and DES(*p*) [(i.e. a desire that *p*)] are the same propositional attitude. The point is rather that DES(*p*) might be equivalent to some other belief—not to BEL(*p*) but to BEL(DES(*p*)). (Price, 1989, 120)<sup>79</sup>

Price goes on to show that we can make the same point in the linguistic case.<sup>80</sup> Indeed

---

<sup>79</sup> Cf. Falk (1947). Price also rejects two other arguments—one by Lewis (1988) and one by Collins (1988). Lewis (1996) continues his support for Humeanism in a slightly different way. For a response to Lewis, see, e.g., Hajek & Petit (2004) and Weintraub (2007). For other responses to M. Smith, see, e.g., Schueler (1991), van Roojen (1995), Sobel & Copp (2001), Coleman (2008), and Alverez (2012, 66 – 71).

<sup>80</sup> (Price, 1989, 121)

Price's point is quite general; things with a dual causal-functional profile are not incoherent at all. To further illustrate, consider Matthew S. Bedke's example of something that has a dual causal-functional profile, namely his leather jacket.<sup>81</sup> Bedke takes it that the jacket serves to keep him warm *and* simultaneously serves to keep him stylish. Our world is rich with things that have these kinds of multiple causal-functional profiles. That something is claimed to possess a dual causal-functional profile provides us with no good reason to reject the claim as incoherent or that accepting it is a theoretical cost.

Hence, understanding Amber's psychological situation as one that recognizes the ill will Benson manifests and, without independent influence, pushes her to rebuke him is not incoherent simply because it does double duty. Likewise, taking Sher's response to his daughter to involve a cognitive judgment that she has manifested a will that is insufficiently good that also pushes him to be more vigilant but no less trusting is not a confused interpretation merely due to the dual function of that judgment.

To further the claim that belief-like states are necessarily distinct or independent from desire-like states, Humeans point out that the former sometimes survive the silencing or defeat of the latter.<sup>82</sup> The objection is that, if the anti-Humean is correct, then one cannot sincerely judge that she morally ought to cease eating meat and fail to be motivated to abide by that judgment. For illustration, we are sometimes referred to Michael Stocker's classic tale:

Recently, I read a story of what might be taken as typical of one course of life. It was said of this political figure that, in his youth, he cared a lot about the suffering

---

<sup>81</sup> Bedke (unpublished ms).

<sup>82</sup> M. Smith (1987; 1988, 594; 1994, 119 – 125).

of people in all parts of the world and devoted himself to making their lives better. But now he concerns himself only with the lives and fortunes of his close family and friends. He remembers his past, and he knows that there is still a lot he could do to help others. But he no longer has any desire so to do. (Stocker, 1979, 741)

The claim, then, is that this *Listless Politician*, as I'll call him, should be *impossible* if anti-Humeanism is correct. The thought is that on anti-Humeanism Listless Politician should be motivated just in virtue of his having a sincere moral belief, but he lacks motivation to help while his sincere moral judgment persists. And, further, the case seems typical not impossible. Thus, Humeans conclude that anti-Humeanism is implausible.

Clearly, this would be deeply troublesome if the anti-Humean were in fact committed to the claim that the belief-like state *entails* the possession of a desire-like, motivation to do as the former prescribes. But, they aren't. Rather, the anti-Humean can hold that sincere moral judgment retains its ability to motivate even where it fails to do so. Like Humeans, anti-Humeans can make use of the explanations that make sense of instances where a belief-desire pair should be motivating an individual to act but isn't.<sup>83</sup> For instance, some other irrationality might be getting in the way.

Generally, then, as Mark van Roojen makes clear, on dispositional accounts of mental states, "a state's direction of fit is its aptness for playing a certain role" (van Roojen, 1995, 47). Thus, van Roojen continues, belief-like states that possess motivational efficacy retain their ability to motivate even where that ability is not "manifested while [an individual is] both in that state, and in generally suitable conditions for the manifestation of its aptness for the role" (Ibid., footnote omitted). So, the "anti-

---

<sup>83</sup> See, Swartzter (unpublished ms).

Humean about motivation [can] deny the charge that the presence of certain cognitive states entails desires, while maintaining that these states have more than one direction of fit” (Ibid., 47). Continuing the leather jacket example for illustration, the point being made is that when in Rome, our leather jacket may cease to be stylish. But, that does not entail that it ceases to keep us warm nor that the jacket lacks the ability to function as a provider of warmth *and* style. Thus, anti-Humeans can make sense of Listless Politician so long as they have a story explaining why his sincere moral belief fails to motivate. Hence, Stocker’s politician may (i) sincerely judge that he has more good to do in the world and (ii) this cognitive attitude may be enough on its own to motivate him to do that good where (iii) given his current depression the manifestation of that motivation is defeated. Likewise, were Amber differently situated, her recognition of Benson’s ill will may have failed to change her in blame-characteristic ways due to some psychologically significant mood, but that does not mean that it is not capable of doing so.

The final concern is (c), i.e. the worry that positing attitudes that are belief-like and desire-like—namely, *besires*—is simply *ad hoc*. Along these lines, Russ Schafer-Landau writes:

One might claim that evaluative beliefs can have two directions of fit, in virtue of having two contents, one aimed at describing the value attaching to some course of action, the other aimed at characterizing the end state being sought. But this seems *ad hoc*, at least if one is willing in all other contexts to distinguish belief from desire in terms of direction of fit. (2003, 173 n. 11)

The idea is that anti-Humeans allow that beliefs and desires can be distinct and they allow that belief-desire pairs where the belief is distinct from the desire can fulfill an explanatory role in the practical realm as well as others. So, the thought continues, it is

curious that anti-Humeans posit these states that have a dual causal-functional profile as well. The objector submits, then, that the appeal to such attitudes is *ad hoc* or unwarranted.

On the contrary, it is not clear that the attitudes appealed to by anti-Humeans lack explanatory value over and above those posited by Humeans. For instance, the Recognition Account has it that such attitudes (a.k.a. besires), as opposed to belief-(frustrated)-desire pairs, facilitate part of the best explanation for the way in which responsibility responses go *beyond belief* as it were in a manner that uniformly captures the emotive-centricity of those responses.<sup>84</sup> So, in this context, citing besires as key to capturing the contours of responding to someone as responsible for her behavior is more than just not *ad hoc*. If the Recognition Account is correct, we have independent reason to acknowledge their existence and the important role they play in our practical lives.

### 3: RECOGNITION AND DESERT

Thus far, the sketch of the Recognition Account illuminates the key functions involved. To more completely illustrate the view, I turn to elucidate the intimacy that obtains between praise-/blame-constituting recognition and the characteristic dispositions of responsibility responses. In particular, I draw attention to one further and important dimension of responsibility, namely, the dimension of desert. In so doing we see that Humean Blame forces a gap between (a) an agent's worthiness of being responded to as responsible and so her worthiness of being the target of the dispositions characteristic of

---

<sup>84</sup> See, e.g., Swartzler (2013) for other aspects of our moral psychology best explained by besires.

such responses and (b) the reason why the respondent is so disposed. Call the gap, *the worthiness gap*.

To see that Humean Blame commits us to the worthiness gap note that responsibility responses can be entirely private insofar as one need not *express* any attitude in order to blame. That view commits proponents to something like:

*Desert:* “X deserves blame” means that blame directed at X is justified or appropriate.<sup>85</sup>

A central worry regarding Desert is that “the normative element of [our favored account of blame] must thus be oriented to the person doing the blaming rather than the person blamed” (Sher, 2006, 86). The concern about blame’s orientation seems to follow from the fact that a commitment to Desert entails adherence to:

*Warrant:* Blame is a response that others are warranted in having when appropriate.<sup>86</sup>

So, on views like the Recognition Account, the concern is that blame is not “something the wrongdoer ‘has coming to him’ but rather [...] a response that his transgression renders appropriate in others” (Ibid., 86). Thus, the concern continues, a view committed to Desert and Warrant “renders mysterious our sense that to call someone blameworthy is to pass a negative moral judgment on *him*” (Ibid., 88, his emphasis). This is at least one way in which such accounts might seem to take “the blame out of blame” (Wallace, 2011, 349).

---

<sup>85</sup> Sher (2006, 86).

<sup>86</sup> Ibid., 86.

On the contrary, a commitment to Desert and Warrant leaves the blame in blame insofar as evaluating an agent as blameworthy is to pass a moral judgment on that agent. Notice, taking up Desert and Warrant is consistent with maintaining that *the reason why* holding a particular individual responsible is appropriate or fitting is that the beholder is simply responding to, say, the wrongdoing that individual has done or the ill will reflected in *her* action. In the end, then, blame involves passing a negative moral judgment on the agent being blamed. If the blamer is correct that someone is blameworthy, then she is correct that *they* have evinced ill will and her blame of them calls attention to that fact.<sup>87</sup>

Still, Sher's way of keeping the blame in blame, as it were, despite a commitment to Desert and Warrant forces a gap between the recognition of blameworthiness and the reasons that support an emotional response to it. To see this, consider first Sher's take on appropriate blame. On Sher's view, the agent's  $\phi$ -ing only directly does two things—namely, it makes true the blame-constituting, inert belief and frustrates the blame-constituting desire. For Sher, two other facts come together to legitimate the blame-characteristic dispositions—namely, (a) the fact that a particular desire was frustrated works in concert with (b) the fact that that desire was normatively appropriate to have by the blamer's lights. On Sher's view, then, it is appropriate for S to blame U for  $\phi$  where (i) it is true that U  $\phi$ -ed, and (ii) the desire that U thwarted by  $\phi$ -ing is supported by the normative theory to which S adheres.

So, suppose Uli stole from Sven just to show that he, Uli, was more cunning than Sven. Imagine also that Sven has a desire that individuals like Uli not steal just to evince

---

<sup>87</sup> Cf. Zimmerman (1988, 38).

superior cunningness and he believes that Uli has done just that. Furthermore, stipulate that Sven's desire is supported by his own normative theory. Finally, then, suppose Sven blames Uli for the trespass. According to Sher, Uli's stealing, which seems to render him blameworthy, only does two things. It makes true Sven's blame-constituting belief and it frustrates Sven's blame-constituting desire. On the Humean view, without the frustration of that desire, Sven's belief provides *no* reason for Sven being blame-characteristically disposed. Furthermore, that the desire was frustrated only provides a reason for Sven's dispositional stance toward Uli due to facts about Sven—namely, that he accepts a particular moral code. Hence, none of the facts that render Sven's blame-characteristic dispositions appropriate has to do (directly) with Uli's being worthy of being the object of those dispositions. Thus, Uli's stealing or that which makes him worthy of blame, only provides Sven with a reason to resent or rebuke Uli in a roundabout way. Thus, nowhere in Sher's picture does an agent's worthiness as an object of blame play a (direct) role in the appropriateness of being disposed to resent or rebuke the blamee. In other words, Sher leaves the blame in blame by basically ignoring blameworthiness where it matters most.

Pushing back, Humeans might point out that Sher is driven to his view by his rejection of the idea that the appropriateness of responsibility responses takes the same basic shape as the appropriateness of attitudes that attempt to represent the world as it is. Indeed, Sher holds that that idea entails that the attitudes constituting a responsibility response of two different people, A and B, toward a single individual, C, could differ with respect to the attitudes' degrees of warrant.<sup>88</sup> Sher thinks that this is unacceptable

---

<sup>88</sup> Sher (2006, 116 – 117).

because it “would imply that an agent can be blameworthy relative to the situations of some people but not blameworthy relative to the situations of others” (Sher, 2006, 116).

Sher’s concern, however, is misplaced. Whether A’s attitude that (partly) constitutes responding to C as normatively responsible for  $\phi$  is *more* warranted or appropriate than B’s attitude does not change the fact that C is worthy of normatively charged responses on the basis of aspects of herself. We must keep distinct (1) the fact that, from an individual’s perspective, C is worthy of being responded to *in a particular way* as responsible for her action in virtue of facts about that individual, C, and the surrounding context and (2) the fact that C is worthy of being responded to as a responsible agent *full stop*. For example, if an act of unjustified betrayal is attributable to Cindy as the basis of negative moral appraisal, Cindy is blameworthy full stop. That is, Cindy gives others cause or reason to blame her whether anyone is in a position to do so or not. Still, given certain relevant aspects of Alexis, Brian, and the context at hand, each may differ with respect to the ways in which it is appropriate for them to respond to Cindy.

The Recognition Account avoids this mistake altogether and leaves the blame in blame, since it is the worthiness of the individual being responded to as responsible that renders instances of the phenomenon appropriate. The idea is that their worthiness is a *pro tanto* reason to respond to them as responsible where that response just is the recognition of that worth or reason and change in/reinforcement of blame-characteristic dispositions. For instance, even if no was around when Amber’s boss, Benson, broke their contract, he was blameworthy insofar as his act provided *pro tanto* reason to blame him. For some, that reason may be completely defeated given facts about them and the

surrounding situation. For others, that reason is only slightly defeated and so renders fitting a way of blaming Benson that is perhaps different from the response it renders appropriate for Amber. Still, it is *that* reason, based in or reducible to Benson's worthiness, that renders blame of Benson appropriate.

Finally, there is a third way of pressing the concern that a commitment to Desert and Warrant does not properly capture the idea that blame is something the wrongdoer 'has coming to him' and so leaves the blame out of blame. This final concern focuses on the fact that it seems as though implicit in the attitudes involved in responsibility responses is some *pro tanto* reason for expressing the attitudes constitutive of that episode.<sup>89</sup> Here, the idea is that the sense in which blame is something the wrongdoer 'has coming to him' is best captured by a thesis like:

*Express:* U is worthy of being responded to as responsible just in case S has *pro tanto* reason to express the attitudes constitutive of S's responding to U as responsible in way W.

The concern, then, is that those committed to Warrant must find room for Express.

Fortunately, proponents of Warrant can adopt Express. We just need to keep separate views about that which bears on the *accuracy* of an attitude and views about that which bears on the *permissibility* of its expression.<sup>90</sup> Note, 'appropriate' as it features in Warrant is concerned with the former, not the latter.<sup>91</sup> Thus, adopting Desert and Warrant says nothing about whether implicit in blame is some *pro tanto* reason for

---

<sup>89</sup> Cf. Pereboom (2013). This is sometimes referred to as *basic desert*.

<sup>90</sup> Cf. Hieronymi (2004). Cf. McKenna (2012, Chpts. 5, 6, & 7, esp. 120 – 140 and 171 – 173) who makes the same mistake. [omitted]

<sup>91</sup> This holds even if something like D'Arms & Jacobson's (forthcoming) view that appropriateness is about the dispositional states that the response sets up is true. But, 'accuracy' may be too cognitive for their liking. Perhaps aptness works better in that context. [omitted]

expressing it. So, allegiance to Express is *not* closed off by adherence to these other theses.<sup>92</sup>

Furthermore, since Humean Blame has it that blame is partly constituted by a normatively legitimate and dissatisfied desire, coupling Humean Blame with Express returns an odd result. The most straightforward coupling of Express with Humean Blame implies that where an individual U is worthy of S's blame there is some *pro tanto* reason for S to tell U in some manner that S believes that U is blameworthy and for voicing the fact that U has dissatisfied one of S's normatively appropriate desires thus disposing S in blame-characteristic ways. This just seems completely inaccurate.

Interestingly, coupling Express with the Recognition Account returns a more natural picture according to which no intermediary talk of desires and their frustration is necessary to fully express the response. Rather, here appropriate blame grounds a *pro tanto* reason for S to in some way tell U that S has evaluated U's behavior as worthy of blame such that S is disposed toward U in a manner characteristic of blame.

The Recognition Account, then, captures a robust sense in which blame is something the wrongdoer 'has coming to him'. Here, the propriety of holding responsible is solely grounded in whether the individual being blamed is worthy of being so held. And, when coupled with the idea that where S appropriately blames U, S has a *pro tanto* reason to express the underlying attitudes, the Recognition Account returns a natural understanding of what S has reason to express.

---

<sup>92</sup> The Recognition Account is also compatible with the idea that responsibility responses implicitly call out for a response. See, e.g., Walker (2003, 157 – 161), Macnamara (2013; forthcoming), and A. Smith (2013).

#### 4: CONCLUSION

Built on the core of the Strawsonian Reactive Attitudes Account, the Recognition Account retains the richness of the former while uniformly making sense of the important place that responsibility responses occupy in our emotional lives even where those responses occur without affect and without aid beyond Reason or the cognitive to play that role. Hence, the Recognition Account offers an attractive alternative to denying affectless blame, digging in with a non-unified view to make sense of the various aspects of responsibility responses, or going in for the less plausible Humean stance. Indeed, as Strawsonians suggest, responding to each other as morally responsible appears to be, at bottom, simply about motivationally efficacious recognition. The Recognition Account simply makes good on that suggestion and illuminates the full contours of our responsibility responses.

Still, some might wonder whether the promiscuity regarding the attitudes that can fulfill the recognition function and the emotion function allowed on the Recognition Account grounds a good reason for resisting it. In the next chapter, I argue that it does not. In the next chapter, I develop the Recognition Account alongside Scanlon's (2008) and Sher's (2006) accounts of blame. Each view shares the blame-constituting promiscuity that some find objectionable. I argue that none of the various ways of pressing the objection are sound.

#### IV: RESTRICTIONS AND RESENTMENT

Sometimes we witness wrongs and recognize them as such. That recognition either transforms or secures the way we interpret our surroundings and our inclinations to react to them. Similarly, the judgment that an individual has done what is right alters or strengthens the judge's interpretational and reactive proclivities. Responding to an individual as normatively responsible, then, involves recognizing a particular quality in the will reflected in their normatively significant act or attitude where that recognition is motivational insofar as it changes or reinforces the responder's attitudinal and behavioral dispositional profile. Views diverge over whether these responses are always, or ever, emotional and felt. Most lean toward always, holding a version of the P. F. Strawson-inspired:

*Reactive Attitudes Account:* A responsibility response is constituted by a complex conative-affective attitude (e.g. a reactive emotion like resentment, indignation, guilt, gratitude, approbation, or pride).<sup>93</sup>

Still, some hold fast to the affectless

*Ledger Account:* A responsibility response is constituted by a belief about an individual as praiseworthy/blameworthy and a positive/negative mark on their ledge of life.<sup>94</sup>

Generally, this essay concentrates on, defends, and develops accounts of blame that fall in between.

---

<sup>93</sup> See, e.g., Wallace (1994), Wertheimer (1998), Hieronymi (2004), and Hurley & Macnamara (2010), McKenna (2012).

<sup>94</sup> See, e.g., Feinberg (1970, 124 – 125), Glover (1970, 64), Zimmerman (1988, 38), and Haji (1998, 8).

On these middle-ground accounts, the psychological state of responding to an individual as morally responsible and blameworthy often, though not always, involves felt affect. For instance, T. M. Scanlon (2008; 2013) rejects the Reactive Attitudes Account but acknowledges that blame sometimes involves blame-characteristic feelings like resentment, indignation, and anger.<sup>95</sup> Likewise, George Sher (2006) adopts an account on which blame is constituted by a particular belief-desire pair that need not bear any felt affect as it disposes the blamer to rebuke or sanction the blamee. And, elsewhere I develop a view on which blame is constituted by a sometimes affectless but nevertheless motivationally efficacious cognitive recognition of ill or insufficiently good will. Though insightful in their own right, middle-ground accounts like Scanlon's, Sher's, and mine suffer from the concern that they are too unrestrictive with respect to the psychological states they allow to count as instances of blame. In this essay, I argue that we have no good reason to think that these accounts draw the boundaries too wide.

To make the case for middle-ground views generally, I'll concentrate on my own

*Recognition Account:* A responsibility response is constituted by an individual S recognizing that an individual U manifested a particular quality of will where that recognition motivates a change in or reinforcement of S's corresponding praise-/blame-characteristic dispositions.

Elaborating a bit, according to my account, responsibility responses are specified by two functions—i.e., the recognition function and the emotion function. The attitudes serving the former recognize a manifestation of an individual's will and its quality. The attitudes

---

<sup>95</sup> Angela M. Smith adopts a view similar to Scanlon's (1998, Chpt. 6). Smith (2005, 250 – 251, esp. n. 26 and n. 27) may, however, be nodding to a middle-ground account about the psychology of blame. Compare that with Smith (2007, esp. 470, 476 – 478) where she distinguishes passive blame from active blame. The latter is emotional and potentially expressive, but the former is not.

serving the latter motivate the alterations or fortifications of the respondent's blame-characteristic dispositional profile in a manner commensurate with facts about the respondent's moral commitments and the content of the recognition. Importantly, my account unlike Sher's is anti-Humean insofar as the cognitive recognition can motivate without the aid of independent desire. Furthermore, emotions and non-emotions can fulfill the functions that specify the state. Thus, some instances of blame involve no felt affect while others do.

The allowance that an agent may blame without felt affect is the target of extensive criticism. Since it is not exactly clear what the common concern is when it comes to the limits an account places on the types of mental states that count as responsibility responses, I address the most well pronounced versions of this concern as expressed by Susan Wolf (2011) (§1), R. Jay Wallace (1994; 2011) (§2), and Victoria McGeer (2013) (§3). I argue that none of these *restriction arguments*, as I call them, successfully shows that the less restrictive middle-ground accounts are any less plausible for being so. I conclude that promiscuity with respect to the attitudes involved in holding responsible offers no reason to reject middle-ground accounts like mine, Scanlon's, or Sher's (§4). I begin with Susan Wolf's restriction argument.

#### 1: WOLF'S RESTRICTION ARGUMENT

According to Susan Wolf, "there is a family of attitudes and activities that are picked out by "blame" as it occurs in [her] dialect, for which it is worthwhile to retain a specific

term” (Wolf, 2011, 335).<sup>96</sup> Wolf offers three reasons in support of her view. First, Wolf cites the positive role that resentment, indignation, anger, and guilt might have for some writing that “it would be a shame to revise our conceptual scheme in a way that minimized their distinctiveness and ignored or denied their value” (Wolf, 2011, 335). Second, Wolf is concerned that some modifications allowed by middle-ground views like the Recognition Account have different and less stringent justifying conditions. Her concern is that we may “fail to distinguish [her] sort of blame from [a less restrictive sort]” and, if we do, “we may not appreciate or attend to these differences, and serious consequences may ensue” (Wolf, 2011, 335). Finally, Wolf thinks that we middle-ground views undermine our comprehension of certain ongoing debates where blame is a central feature. In fact, she thinks that talking in terms of these middle-ground accounts distorts what is at issue in those debates. Wolf writes:

We cannot understand the history of the free will debate without making reference to this set of attitudes [...], nor can we do justice to the continued discussion of the problem if we fail to recognize that the intelligibility and legitimacy of this set of attitudes [...] in particular is at least part of what is at stake” (Wolf, 2011, 335).

Wolf concludes, then, that middle-ground views like the Recognition Account should be rejected, since they fail on these grounds.

Wolf’s rejection is premature.<sup>97</sup> First, I fail to see why the less restrictive

---

<sup>96</sup> Susan Wolf (2011) is directing her comments at T. M. Scanlon’s account of blame. However, her concern generalizes to other middle-ground views like Sher’s and mine.

<sup>97</sup> It should be noted, perhaps, that Wolf in some ways seems to focus on the *meaning* of ‘blame’. But, neither Scanlon’s account, Sher’s account, nor my Recognition Account is a view about the meaning of terms. Indeed, Sher is quite explicit on this point (2006, 112 – 113). Compare Scanlon (2008, 151). Rather, these views including the Recognition Account are accounts of a particular phenomenon. Very roughly, the phenomenon that these accounts attempt to capture is, a psychological reaction accompanied or constituted (in part) by deep normative or moral appraisal of oneself and others. When this appraisal is

accounts require us to change our conceptual scheme in a way that minimizes the distinctiveness of certain reactive attitudes and ignores or denies their value. As Coleen Macnamara (2011, 101) makes clear, even if you think restricting the class of attitudes that count as instances of holding responsible is mistaken, you could still think that the attitudes like those to which Wolf draws our attention serve a distinctive function and have a particular value.<sup>98</sup> The Recognition Account and other middle-ground views are simply (generally) silent about what that function or value is.

Second, it is likewise unclear why we should think that middle-ground accounts render it likely that we will fail to recognize differences pertaining to appropriateness conditions. On the contrary, on the Recognition Account the only attitudes that feature in responsibility responses are those that, like the reactive attitudes, are sensitive to and represent qualities of will. Likewise, the cognitive components on Scanlon's and Sher's view are responsive to normatively significant reasons like flouting a moral norm, doing something wrong, or possessing a vicious character trait.<sup>99</sup> So, despite the fact that middle-ground accounts allow that blame may involve more than just the sentiments

---

negative, it is generally referred to as 'blame', and when positive, 'praise'. Still, the phenomenon might be broader than the applicability conditions of either term allow. So, whether using the terms 'blame' and 'praise' as shorthand for that broader phenomenon is legitimate is beside the point. Indeed, agreeing with Wolf that the meaning of 'blame' is not captured by the Recognition Account is compatible with accepting it as an account of the relevant phenomenon. Thus, Wolf's criticisms as focused on the meaning of 'blame' seem out of place here. Furthermore, even if Wolf's critique were on target, it is generally mistaken in the ways I mark in the text.

<sup>98</sup> Macnamara is focused on Wallace's (1994) account. There the only attitudes that count as instances of blame are resentment, indignation, and, for self-blame, guilt. To be clear, I disagree with the details of Macnamara's exact diagnosis of Wallace, which sees him as subtly shifting between two different senses of 'holding responsible'. Wallace is clearly aware of these as evidenced by Wallace's (1994, 23) clarificatory statements right before he moves to narrow the class of blame-constituting reactive attitudes.

<sup>99</sup> See, e.g., Scanlon (2008, 141 – 143) and Sher (2006, 115 – 119). It is unclear how to translate Scanlon's use of 'significance'. Though his previous sympathies (e.g., Scanlon 1988) lie in a quality of will or quality of decision sort of view, this may not helpfully capture his new position.

upon which Wolf focuses, we have no good reason to suppose that such promiscuity entails differences regarding their appropriateness conditions.<sup>100</sup>

Finally, I also fail to why Wolf thinks that we are unable to understand or do justice to the free will debate concerning moral responsibility. At least one thought leaves me in doubt.<sup>101</sup> Note, even if Wolf is correct that the only way to properly understand or do justice to the debate is by thinking of it solely in terms of her restricted class, all that follows is that seriously engaging in that debate requires restricting ourselves to discussing those attitudes. So, for instance, perhaps just showing that disappointment is justified in the face of determinism is insufficient to shore up support for compatibilism. On this supposition, one would have to show that resentment, indignation, anger, and guilt are justified in the face of causal determinism or some such. But, when simply attempting to capture the phenomenon in question, it would be a mistake to observe such a restriction on Wolf's grounds. Note, in general, just because seriously engaging a debate featuring a particular phenomenon requires that we restrict conversation to certain aspects of it does not entail that the phenomenon is wholly comprised of just those aspects.

---

<sup>100</sup> In case I am being somewhat opaque here, the basic idea is that at least one important axis of appropriateness simply concerns whether the appraisal embodied in the episode of blame is accurate or correct. I take it that the accuracy or correctness depends on whether the will taken to have been manifested was in fact manifested and whether it has the quality the appraisal claims it has. Thus, two states that track manifestations of regard are the same insofar as this axis of appropriateness is concerned. [omitted]

<sup>101</sup> A second is that, as far as I can tell, we have no reason to believe that the debate *does* concern *only* resentment, indignation, anger, and guilt. Even if those instances of responsibility responses are presumed to be the most salient due to the fact, if it is a fact, that they are often accompanied by a certain degree of rancor, it seems that we might follow Strawson more closely and take it to be a debate about all of the episodes that occur in recognition of qualities of will.

Moreover, Wolf's favored set of attitudes have whatever distinctness and value they may have, carry the appropriateness conditions they may carry, and are central to the free will debate only in virtue of what they are with respect to responding to an individual as responsible. The Recognition Account identifies them, in this respect, as modifications in recognition of qualities of will that affect changes in or reinforce our relevant behavioral and attitudinal dispositions. Wolf's critique fails to show that the Recognition Account rules out correct views about the distinctness and value of the various attitudes involved, their appropriateness conditions, and/or their centrality to particular debates. Thus, Wolf's critique does not show against the recognition account. Much of the same could be said in response to Wallace (2011).

## 2: WALLACE'S RESTRICTION ARGUMENT

Recently, R. Jay Wallace argued that accounts of responsibility responses on which attitudes other than the reactive attitudes can feature "leave the blame out of blame" (Wallace, 2011, 355). Wallace relies here on his earlier claim (1994, 25 – 40) that restricting the class of attitudes involved in blame to just those of resentment, indignation, and guilt puts us in a position to make significant moral distinctions.<sup>102</sup> Wallace's new argument, then, is only as strong as his earlier argument, since he is here simply attempting to capitalize on his earlier view by focusing on what he sees as the *opprobrium* distinctive of blame.<sup>103</sup> In favor of his restriction, Wallace basically claims

---

<sup>102</sup> Wallace (1994, 25 – 40). The only added nuance to Wallace (2011) is that he attempts to put weight on what he calls the *opprobrium* that comes with the blame-constituting reactive attitudes.

<sup>103</sup> Wallace (2011, esp. 348 – 349).

that it “can help illuminate our moral and psychological categories” (Wallace, 1994, 33). According to Wallace, this illumination comes in four flavors. First, Wallace argues that his restriction allows us to sharply distinguish blame-constituting emotions from other moral sentiments like shame, contempt, and derision. Second, Wallace argues that his restriction allows for a clear distinction between the *moral* and *non-moral* attitudes involved in responding to an individual as morally responsible or blameworthy. Third, Wallace argues that restricting blame-constituting attitudes to resentment, indignation, and guilt supports the Strawsonian idea that there is a strong opposition between the reactive attitudes and the objective stance or attitude—i.e., the broad psychological perspective that colors individuals as natural, mechanical things. Finally, Wallace contends that his restriction allows us to distinguish between the natural emotions and the reactive emotions, something Wallace thinks Strawson failed to do because he did not restrict the class of blame-constituting attitudes in the way that Wallace does. If Wallace is correct that his restriction alone allows for these four illuminations of our moral psychology, then we have strong reason against Wallace’s target, Scanlon’s (2008) view. But, his criticism would also extend beyond Scanlon to all cause serious trouble for middle-ground views generally.

*Pace* Wallace, however, his restriction alone is not what does the work here. In fact, the restriction itself seems to play no role at all. As I show presently, the heavy lifting is done by the character traits Wallace claims the emotions in his restricted class have. But, even if they have those traits, Wallace gives us no reason to believe that other attitudes have those traits as well.

To see this, consider Wallace's claim that his restriction illuminates the distinction between the blame-constituting moral emotions and other moral sentiments. For Wallace the reactive attitudes are distinct from other moral sentiments in virtue of the fact (if it is a fact) that the former are sensitive to instances of right and wrong while the latter aren't.<sup>104</sup> On Wallace's view, the other moral sentiments are sensitive to "the various modalities of moral value, such as the values of kindness or consideration or benevolence or even justice" (Wallace, 1994, 37).<sup>105</sup> In other words, blame-constituting moral emotions unlike other moral emotions recognize compliance "with moral principles [which] makes it possible for us to relate to people on terms of mutual recognition and regard" (Wallace, 2011, 369). But, on its own that appeal does not establish that any *particular* attitude should be included or excluded from the set of blame-constituting attitudes. Furthermore, it's not that we would seek to include shame, derision, or contempt in the class of blame-constituting attitudes. Rather, there may be other attitudes that recognize compliance with moral principles.

Further still, Wallace give us no reason to believe that blame is always a recognition of such compliance. On Scanlon's view, blame is generally (or at least also) a response to the significance that an individual's attitudes and behaviors have for our interpersonal relationships. And, on Sher's view, blame is sometimes directed at character traits. On that construal, it does not seem that blame is recognizing compliance with moral principles, at least not present compliance. Finally, the Recognition Account, need not adopt a construal of blame on which blame essentially involves a recognition of such compliance. Certainly, the Recognition Account has it that blame essentially

---

<sup>104</sup> Wallace (1994, 37 – 38). See also Wallace (1994, 127 – 136).

<sup>105</sup> Cf. Wallace (2011, 366 – 369).

involves the recognition of ill or insufficiently good will. But, as David Shoemaker's (forthcoming a, b) makes clear, there are equally plausible readings of what quality of will comes to such that the recognition I argue is central to blame is wider than simple compliance with moral norms and principles.

Therefore, middle-ground views can accept the distinction Wallace is driving at without agreeing that the only attitudes directly involved in cases of blame are Wallace's favored three. So long as these accounts allow that the blame-constituting attitudes are essentially responsive to something that the non-blame-constituting attitudes are not, the distinctions can be easily preserved. And, it isn't clear at all that we should agree with Wallace that the blame-constituting attitudes are only responsive to the failure to comply with moral principles. Nor, is it clear why we should think that attitudes beyond Wallace's favor are not responsive to that failure. that which reflects the extent to which individuals adhere to moral norms the account can make the important distinctions to which Wallace draws our attention. And, moreover, And, since (a) Wallace gives us no reason to think that only emotions of a certain sort can do that and (b) the Recognition Account like Scanlon's and Sher's agrees with Wallace that the attitudes involved in a case of blame are sensitive to indications of departures from moral norms, the less restrictive middle-ground accounts are no less able to help illuminate our moral and psychological categories. Wallace's oversight repeats.

For instance, take the second distinction Wallace thinks his restriction illuminates—namely, the difference between *moral* and *non-moral* attitudes involved in responding to an individual as morally responsible or blameworthy. Wallace contends that both moral and non-moral *reasons* are represented in behaviors that evince an

agent's quality of will.<sup>106</sup> However, on his view the *moral* attitudes track or are sensitive to violations of obligations supported by *moral* reasons. Alternatively, *non-moral* reactive attitudes track or are sensitive to violations of obligations supported by *non-moral* reasons. Hence, Wallace thinks that the former but not the latter “have a distinctively moral content” (Wallace, 1994, 36) and the only moral reactive attitudes are his favored three—resentment, indignation, and guilt.

Whether Wallace is correct or not to distinguish moral attitudes from non-moral attitudes in the way he does, it should be clear that nothing about the restrictiveness of Wallace's class helps to distinguish them. In fact, Wallace offers no argument for the claim that only resentment, indignation, and guilt are sensitive to moral reasons. Nor does he offer an argument for the claim that the reactive attitudes are only responsive to violations of obligations. Thus, so long as proponents of middle-ground accounts (a) distinguish between moral and non-moral reasons and (b) hold that attitudes that are sensitive to violations of obligations supported by moral reasons are the only ones with distinctively moral content, we can draw the same distinction without adopting Wallace's restricted class. Furthermore, if we have good reason to believe that moral reasons underwrite other norms, e.g., norms regarding the proper character to cultivate, then we replace (b) with a suitable alternative and still make the distinction Wallace want to make. .

Consider also two further ways in which Wallace (1994) supports his restricting blame-constituting reactive attitudes. Each appeals again to the idea that the restriction

---

<sup>106</sup> Should we adjust this to say that the moral reactive attitudes track morally significant qualities of will whereas the non-moral reactive attitudes track morally insignificant qualities of will or some such? Perhaps, but Wallace is fairly unclear here. Moreover, nothing here turns on employing my distinction or this adjustment.

sharpens moral and psychological categories. However, this set of reasons is narrowly focused insofar as it is aimed at the development of Strawson's view. Here, Wallace is attempting to show that with his restriction in place, we can do better than Strawson did at illuminating certain aspects of our moral psychology.

For example, Wallace claims that his "narrow construal of reactive attitudes can be taken to support two further and more important claims that Strawson wishes to defend" (Wallace, 1994, 30). The first "is Strawson's claim that the reactive attitudes are opposed to the objective stance of the scientist or therapist or the executor of social policy" (Ibid., 30). And, the second claim is that individuals are naturally susceptible to the attitudes involved in holding an individual responsible. Wallace thinks that his restricted class can capture both claims "without simply conflating the natural and reactive emotions, as Strawson seemed to do" (Ibid., 32).

On the contrary, Wallace's narrow class of blame-constituting reactive attitudes is not what does the work here. Seeing this requires piecing some important bits of Wallace's earlier book together. Note, first, that Wallace tells us:

...the moral reactive emotions are only fair responses when directed toward people whom we see as potentially standing in a certain kind of moral relationship with us, one distinguished by the exchange of moral criticism and justification [...and...] this is not the way in which therapists or scientists conceive their professional relations with their subjects. (Ibid., 31 n. 17)

Second, Wallace says that he "will not be in a position to explain this phenomenon until the argument in Chapters 5 and 6 is complete" (Ibid., 31 n. 17). After piecing together Wallace's argument, I show that the less restrictive middle-ground accounts can shed just as much light on those aspects of our moral psychology that Wallace claims are central

but murkier in Strawson's original. As we'll see, Wallace's argument only supports the view that capturing Strawson's thoughts requires accepting either a control condition on appropriate blame or a sociological claim about scientists and therapists. Middle-ground accounts can adopt either or both.

Curiously, despite claiming that he will do so Wallace never explicitly explains the phenomenon he refers to in the above quotations after completing his arguments in Chapters 5 and 6. Still, we can gather an explanation from the main themes of those chapters. Primarily, Chapters 5 and 6 seek to support a view about when, if ever, blame is appropriate.<sup>107</sup> For Wallace, generally, holding responsible is appropriate when doing so is fair.<sup>108</sup> Further, Wallace identifies what he takes to be excuses and exemptions each of which undermine the fairness or appropriateness of blame. Aligned generally with Strawson, on Wallace's view, an individual is excused or exempt in virtue of the fact that a particular condition makes it so that their behavior evinces no will good or ill.<sup>109</sup> According to Wallace, the disposition to respond to an individual as responsible via attitudes like resentment, indignation, and guilt are tied to moral expectations and demands—expectations of and demands that one be shown a particular degree of regard.<sup>110</sup> And, he holds that such regard is only demonstrated by those with *reflective self-control* and in the absence of excusing conditions.<sup>111</sup>

I take it, then, that there are two salient and reasonable ways to interpret his point

---

<sup>107</sup> Wallace (1994, 115, 118, and 154 – 155).

<sup>108</sup> Wallace (1994, Chpt. 5, esp. 118 – 136, and Chpt. 6, esp. 154 – 166).

<sup>109</sup> Wallace (1994, Chpt. 5, esp. 118 – 136, and Chpt. 6, esp. 154 – 166).

<sup>110</sup> Wallace (1994, 127 – 136).

<sup>111</sup> Wallace (1994, 194). Also, this last bit I have worded in a way that almost deserves quotation marks. Given the structure of the sentence, the marks would have come in places that reduce clarity.

regarding the objective stance.<sup>112</sup> On the one hand, Wallace might think that therapists do not see their patients as individuals possessing the control requisite for violating moral expectations. On the other hand, Wallace might think that these individuals (i.e. therapists and scientists) simply do not concern themselves with the regard manifested by the decisions of their patients *qua* patients. Either way, Wallace is in a position to draw the conclusion that he thinks Strawson wants to draw—namely, that therapists and scientists do not react to their patients’ failings with the attitudes operative in blame. *Nevertheless*, that conclusion does *not* rely on the truth of Wallace’s claim that the only blame-constituting attitudes are resentment, indignation, and guilt. Rather, the conclusion about therapists and their patients either depends on (a) the truth of Wallace’s view that control is required in order that the decisions manifest a quality of will *or* it depends on (b) a sociological claim about how therapists see and interpret the behavior of their patients. There is nothing barring middle-ground accounts from adopting either (a) or (b). Thus, restricting the class along Wallace’s lines does nothing to support or better capture the Strawson’s moral psychology as he understands it.<sup>113</sup>

---

<sup>112</sup> For perspicuity, it should be pointed out that one might reasonably hold that the real upshot of these chapters is that we have a viable alternative to those who hold that one is only responsible if she possessed libertarian free will of some sort. Support for this claim might come by way of citing the fact that Wallace rejects the incompatibilist understanding of when and why responsibility is appropriate or fair (Wallace, 1994, 105 – 109). The point is well taken, but then I just do not understand how completing his arguments in Chapters 5 and 6 help us see how his restricted class supports the Strawsonian view that the reactive attitudes are incompatible with the objective stance. Moreover, what I mark as the upshot is what does the work for Wallace in his rejection of the incompatibilist understanding of when and why responsibility is appropriate or fair. So, whether this is an alternative way of understanding Wallace does not affect my arguments here.

<sup>113</sup> Furthermore, I seriously doubt his sociological claim about therapists and scientists. My doubt is in large part due to the extensive literature on the doctor-patient relationship detailing the special sorts of obligations thought to arise from that relationship not to mention scores of health care professionals who, for example, deny women induced abortions or ‘fire’ patients who refuse to follow the doctor’s prescribed

Wallace's second attempt to show that his restriction on the blame-constituting reactive attitudes elucidates distinctions in moral psychology that align with Strawson's own thoughts on the matter also fails. Here, Wallace claims that his narrow class allows us to argue for the view that individuals are naturally susceptible to the attitudes involved in responding to an individual as responsible "without simply conflating the natural and reactive emotions, as Strawson seemed to do" (Wallace, 1994, 32). But, Wallace fails to notice that this has *nothing* to do with the fact that the members of his class are resentment, indignation, and guilt. Rather, it simply has to do with the fact, if it is a fact, that the attitudes involved in responsibility responses are connected to expectations and demands insofar as they are sensitive to or recognize the quality of an individual's will manifested, say, where she violates one of those expectations.

For Wallace, recall that expectations are *moral* where they are supported by moral reasons, which renders them *moral* obligations in the eyes of those who accept them.<sup>114</sup> Further, Wallace treats "moral obligations as governing the choices expressed in action, for it is these *qualities of will* that are subject to the direct influence of reasons" (Wallace, 1994, 157, my emphasis).<sup>115</sup> Thus, Wallace has it that, the expectations that render us

---

treatment plan citing expectations and obligations as grounds for denying them the care they seek. Admittedly the concrete cases of health care professionals 'firing' patients or refusing treatment are subject to various interpretations. I think, however, that a strong case could be made that these professionals by and large see their patients in a light Wallace denies they do. It is beyond the scope of this paper to offer the argument in favor of my interpretation. So, I place very little weight on this as a reason to doubt the veracity of Wallace's claim. Also, we should be especially careful about the emotion work we take to be a part of our vocational lives, especially on the basis of unsubstantiated empirical claims championed as a fact about our moral lives. At the very least, we run the risk of unjustifiably seeing certain vocations as involving no emotional work at all, which in turn may make us unjustifiably unsympathetic to complaints that one's vocation alienates them from themselves. On this latter topic, see, e.g. Talbot Brewer (2011).

<sup>114</sup> Wallace (1994, 36 & 127 – 136).

<sup>115</sup> According to Wallace, "holding someone morally responsible for a particular act presupposes a state of

susceptible to the attitudes involved in blame, concern qualities of will. And, that is just to say that these attitudes are sensitive to manifestations of qualities of will. So, in the end, the restrictiveness of Wallace's set of attitudes involved in blame does no heavy lifting. Rather, we manage not to conflate the *moral* attitudes operative in blame with other attitudes by simply identifying what they are sensitive to, what they respond to, or what they recognize and seeing that non-moral attitudes are not likewise sensitive to such things.

Therefore, Wallace like Wolf fails to offer any good reason to restrict the class of attitudes that may feature in responsibility responses like blame. Hence, he fails to give us a good reason to reject middle-ground views like the Recognition Account on grounds of promiscuity. Perhaps Victoria McGeer can do better.

### 3: MCGEER'S RESTRICTION ARGUMENT

Like Wolf and Wallace, Victoria McGeer (2013) argues that certain emotions are the only attitudes that feature in episodes of blame. However, for McGeer, blame is a species of anger, thus she allows that the class of blame-constituting reactive attitudes may be wider than Wolf's or Wallace's. McGeer's view runs close to mine insofar she holds that blame is a state specified by the function it serves. Further, her view appears to be translatable into a story about responsibility responses generally. Our accounts differ, though, in at least one very important way. On the Recognition Account as with other

---

belief, namely the belief that the act violates a moral obligation one accepts" (Wallace, 1994, 133), and, so, "determines the propositional object of a particular state of reactive emotion such as resentment or indignation, fixing what it is that the resentment or indignation is *about*" (Wallace, 1994, 133, his emphasis).

middle-ground views, blame-constituting attitudes are not essentially a species of anger. In fact, these attitudes can even fail to be emotions. It is this point of difference that McGeer would press that middle-ground views are too unrestrictive or in her terminology they are guilty of attempting to *civilize* blame.

More precisely, McGeer argues that middle-ground accounts are guilty of civilizing the uncivilized and so miss the target phenomenon. McGeer writes:

...civilizing blame amounts to sanitizing it—in effect, to purifying blame of what they [proponents of middle-ground views] take to be features they find normatively problematic; thus, by my lights, it is a psychologically unrealistic theory. Defenders of the sanitizing approach think, further, that this psychologically purified version of blame is normatively satisfying so far as certain features are eliminated. [...] In rushing to reform the psychology of blame, they actually misrepresent what is normatively problematic about those features and so end up with an account of blame that fails to meet either psychological or normative desiderata. (McGeer, 2013, 163 – 164)

On McGeer's view blame must be understood as “emotionally toned, manifested in a variety of expressive behaviors that will often be experienced as punitive by those to whom it's directed” (Ibid., 166). Thus, for McGeer any justification of blame as a normatively appropriate response to the behavior of others must acknowledge the emotionally toned and punitive aspects of it. McGeer objects that middle-ground accounts fail to acknowledge blame's emotional intonation and/or punitive force. Similar to Wolf, then, McGeer holds that where proponents of a view like the Recognition Account offers a justification of blame, they will likely fail to justify the sort of response that often features in our moral and emotional lives.

The Recognition Account escapes McGeer's *civilizing-blame objection* as do Scanlon's and Sher's middle-ground accounts. In short, they escape on at least two

grounds. First, they are not built on facts about the general “normative acceptability” of blame. Whether blame is, in the end, acceptable is a further question. Still, Scanlon and Sher seem to take as given the normative acceptability of all instances of blame from the meek and mild to the fiery and furious. What they seek it seems is an explanation of this normative acceptability. Part of this requires adopting a view about the nature of the phenomenon in question. Like many Strawsonians who favor the Reactive Attitudes Account, Scanlon sees blame as an important indication of the moral or practical health of our interpersonal relationships. And, Sher sees it as a necessary component to the reality of morality. True, Scanlon, Sher, and I reject the idea that blame is essentially constituted by emotions that are basically kinds of anger as well as the idea that blame is always a kind of sanction. But, given that blame as we understand it is central to interpersonal relationships, it should be clear that all sorts of blame play important roles in our practical lives albeit those roles are typically marked out for the reactive emotions to fulfill. As such, even our middle-ground views hold that when blame is expressed, it will likely feel punitive.

Furthermore, middle-ground accounts can also acknowledge that whether blame is embodied solely in reactive emotions or in other attitudes it can be punitive even when left unexpressed. The core weight of blame is contained in what it indicates about the individual who is its object. Thus, the central component to blame’s force is its content.

As Pamela Hieronymi tells us:

...the content of a judgment of ill will can carry a certain amount of force [...] If it is true, then you no longer stand in such a relationship. [...] It also seems quite plausible that standing in relations in which the quality of one’s will is recognized, both by oneself and by others, is of considerable importance. A change in what you or another person thinks about the quality of your will, in

itself, changes your relations with them. Insofar as it is important to stand in relations in which goodwill is recognized, the judgment that you have shown ill will itself carries a certain force. That judgment—even if incorrect—makes it the case that you no longer stand in relations in which your good will is recognized on all sides [i.e. relations of mutual regard]. (2004, 124)

Since the force of blame resides primarily in the content of the recognition that one has failed some normative standard and that recognition is retained on middle-ground accounts as partly constitutive of blame, McGeer is simply mistaken that these latter accounts drain blame of its unwelcome nature.

Moreover, McGeer's charge that middle-ground accounts rush to sanitize blame and thereby distort the psychology involved is misplaced. Scanlon, Sher, and I certainly depart from the Reactive Attitudes Account. However, our departure is not founded on the view that blame is emotionless or only involves fairly mundane emotions. Rather, it is founded on the idea that the phenomenon cannot be wholly accounted for by appealing to a special class of emotions like the reactive attitudes or even anger like sentiments. The accounts are then constructed against a background that attempts to acknowledge the reactive emotions as paradigms of a particular and central aspect of blame. For Scanlon, the search for a better account starts by relocating the emphasis from the moral reactive attitudes to "the expectations, intentions, and other attitudes that constitute these [human] relationships" (Scanlon, 2008, 128). On Sher's view, a key desiderata for an adequate account is that it captures "the role that blame actually plays in our emotional lives" (Sher, 2006, 77). Sher elaborates:

That blame plays some significant emotional role can hardly be denied. This is evident from the energy we expend in trying to affix it, from the rancor that often accompanies it, and from the urgency with which we seek to avoid it. Whatever

their errors, those who view blame exclusively as a mechanism of social control are clearly right to hold that being blamed is often unpleasant enough to have deterrent effect. Thus, if the view that blame is a belief about a person's moral balance or the state of his character is to be at all tenable, its proponents must somehow explain why this is so. (Ibid.)

Similarly, in rebooting the Reactive Attitudes Account, I took Sher's charge quite seriously. The Recognition Account aims to incorporate blame's characteristic emotional import while preserving the beauty of the Strawsonian view. On the Recognition Account, the blame-constituting recognition is (often) itself motivationally efficacious. And, like emotions, that recognition, then, can alter or further secure the blamer's blame-characteristic psychological elements like her interpretive proclivities, her inclinations to withdraw or act out, and her disposition to believe or expect things about and from her world.<sup>116</sup> Thus, each of these middle-ground views seeks to accommodate blame's emotional import and punitive feel when expressed not diminish it. Nor were those accounts designed by cataloguing aspects of blame that make it undesirable. Rather, Scanlon, Sher, and myself seek to better understand blame's undesirable aspects so as to offer a realistic picture of it.

But, there is more to McGeer's view. Consider, for instance, that on McGeer's view, the attitudes involved in blame have backward-looking *and* forward-looking aspects. For McGeer, the backward-looking aspect appraises the harm done while the forward-looking aspect basically seeks to communicate with the transgressor in a manner that aims to cajole her into conforming to the norms she flouted or ignored in the

---

<sup>116</sup> My original inspiration for this thought was Hurley & Macnamara's (2010) account of what it is to have an emotion. [omitted]

future.<sup>117</sup> Some might worry that unless these additions are incorporated proponents of middle-ground accounts will fail to adequately meet McGeer’s conditions on an acceptable account of blame.

To further complicate matters, adopting the view that blame involves a forward-looking aspect might be objectionable in itself. And, McGeer worries that some will reject her view since she holds that responsibility responses are both backward- and forward-looking and, hence, she seems to run counter to

*Anti-Managerialism*: Responding to an agent as responsible is not merely a matter of managing the decisions and actions of that individual in a manner that herds or corals them or aims at such.<sup>118</sup>

McGeer’s concern is that her view is objectionable if it turns blame into “an essentially conditioning enterprise in which the blamer fails to treat the one who is blamed with appropriate moral respect (i.e. fails to acknowledge his or her moral personhood appropriately)” (McGeer, 2013, 176).

A new sort of concern arises for middle-ground views when we take (a) the idea that blame is both backward-looking and forward-looking together with (b) the point that we must accommodate *Anti-Managerialism*. This new concern can be expressed best as a dilemma: either adopt the view that the attitudes involved in responsibility responses have a forward-looking aspect or fail to fully answer McGeer’s restriction argument.<sup>119</sup> But, the thought continues, if we adopt McGeer’s view about the dual nature of the

---

<sup>117</sup> McGeer (2013, 171 – 172) offers empirical reasons for adopting the view that blame has a forward-looking and backward-looking aspect.

<sup>118</sup> *Ibid.*, 176.

<sup>119</sup> [omitted]

attitudes involved in blame, then middle-ground accounts run afoul of Anti-Managerialism.

Fortunately, McGeer persuasively defends against the complaint that her view is contrary to Anti-Managerialism. Furthermore, her defense is one that proponents of middle-ground accounts can accept. To see this, consider the two ways in which McGeer responds to the concern.

First, McGeer points out that the complaint rests on a mistake—namely, the mistake of collapsing the “functional rationale” of a practice with that which “we explicitly aim at when we engage in the practice” (McGeer, 2013, 177). She writes:

Fulfilling that function may be an essential and, indeed, explicitly recognized by-product of the practice in question; yet it is not, perhaps cannot be what practitioners aim at without spoiling the functional efficacy of the practice in question. This familiar observation is what lies behind the hedonistic paradox: happiness is not something we can or do achieve by directly seeking it; yet we often engage in particular practices that aim at something else because we recognize that they make us happy, and we would give them up if they did not. (Ibid., footnote omitted)

Thus, McGeer holds that regulation is simply a by-product of blame. Following McGeer, call this the *Essential By-Product Response*.<sup>120</sup>

McGeer’s second response to the worry that her view is out of step with Anti-Managerialism is that the concern overlooks the fact that intending to change one’s attitudes or behaviors is consistent with “fully respecting them as persons—as responsible moral agents in their own right” (Ibid., 178). So, rather than simply slap a close friend who has wronged him, Hugh might first abnegate them from his life for a

---

<sup>120</sup> Ibid., 177.

short time before convening with them over coffee. Call this the *Regulation-Means Response*.

I take it that the Essential By-Product Response and the Regulation-Means Response are basically correct and clearly compatible with the middle-ground accounts. Thus, whether these accounts agree that blame has forward-looking aspects, they can avoid the concern that that agreement commits them to a view that rejects Anti-Managerialism. Therefore, the proposed dilemma is not vicious.

Neither McGeer's original civilizing-blame objection nor the concerns that stem from Anti-Managerialism offer reasons to reject middle-ground accounts like Scanlon's, Sher's, or mine. Thus, like Wolf and Wallace, McGeer offers no good reason to adopt a more restrictive account of blame.

#### 4: CONCLUSION

Like the Reactive Attitudes Account, the middle-ground accounts of Scanlon, Sher, and myself hold that our stance toward others as participants in a moral community structures the attitudes involved in responsibility responses. Hence, our views paint blame-constituting attitudes as attitudes grounded in the fact that we feel, respond to, and demand of ourselves and others for ourselves and for others as persons worthy of a certain modicum of respect and regard. So, when Jack blames Lynn for wrecking their friendship, we correctly understand his attitude as one that recognizes and calls attention to his personhood and enjoins Lynn to see him as such and respond appropriately. Where expressed, Lynn will likely be upset. Jack can choose expressions that lessen the blow,

but he cannot change what it says about their friendship without failing to express his concern.

The same is true where Dario stands up to his fellow shift manager who is berating a clerk. Like the Reactive Attitudes Account, middle-ground accounts see Dario as expressing (at least in part) his recognition that Luz is treating the clerk in a manner that is objectionable from, say, the standpoint of decency. Dario's demand that Luz take a walk to cool down shows that he acknowledges the personhood of the clerk. It can also encourage everyone to keep this in mind as they attempt to make it through their work day. Further, it speaks to the solidarity we should have for each other as persons in complex interpersonal relationships be it manager to crewperson, queen to serf, or brother to stranger.

Therefore, though it is true that middle-ground accounts are less restrictive about which attitudes are blame-constituting, they maintain the shine that draws us to the Strawsonian view. As Wolf, Wallace, and McGeer's restriction arguments reveal, a key part of that shine is due to the Reactive Attitudes Account's way of making sense of the practical import, the emotional import, and the force of the attitudes central to our moral responsibility concepts, attitudes, and practices. What they fail to reveal, however, is any reason against promiscuity about blame-constituting attitudes. In fact, it seems promiscuity here, perhaps as it is elsewhere, might be preferable.

Still, middle-ground accounts might be taken to civilize blame due to what some see as a worrisome superficiality in the appraisals they license. In the next chapter I address this concern and further develop the Recognition Account as a Real Self View or, as I call it, a *Self-Revelation Account* of moral responsibility.

## V: SELF-REVELATION AND RESENTMENT

Often an individual's behavior reveals who she is as a practical agent and so can serve as the basis of moral appraisal. In other words, behavior tied to an individual's agential activity is (the only) behavior for which she is morally responsible. This paper develops and defends that unified, general conception of responsibility which is codified by

*Appraisal-Basis Thesis:* An individual, S, is responsible for  $\phi$  if and only if  $\phi$  can be attributed to S as the basis of normative appraisal.<sup>121</sup>

*Self-Revelation Thesis:*  $\phi$  is attributable to an individual, S, as the basis of normative appraisal if and only if  $\phi$  reveals who S is as a practical agent.

A key part of my contribution is to bring out, shape, and employ

*Agential Self-Revelation:* An individual, S, reveals who she is as a practical agent via  $\phi$  if and only if  $\phi$  is the result of, is sustained by, or is directly under the control of her agential activity.

Call this picture of responsibility, *the Self-Revelation Picture*.<sup>122</sup> My development of this picture runs along two axes. Each axis is grounded in a concern about whether the Self-Revelation Picture can make sense of the multi-dimensionality of moral appraisal.

---

<sup>121</sup> See, Scanlon (1998, 250; 2008, 202) & Smith (2008; 2012). Cf. Fischer (1987, 63).

<sup>122</sup> Referred to also as *real-self views* (Benson (1987) & Wolf (1990, Ch. 2)) and *self-disclosure views* (Watson (1996, 261 – 262) & Smith (2008)). I opt for 'reveals' to avoid certain implications. For instance, disclosure suggests that an agent *intentionally* brings to light some aspect of her agency. It also suggests that an audience's failure to grasp what is revealed entails a failure of successful disclosure. [omitted]

One dimension of moral appraisal is its depth. Moral appraisal seems to highlight and assess something for which the individual is at fault and so is not superficial and does more than merely describe a person's characteristics. The depth of moral appraisal, then, is the first axis of development. I use as a backdrop Angela M. Smith's (2008) attempt to capture moral appraisal's depth for the Self-Revelation Picture. Smith's attempt appeals to the idea that being responsible for something is to be answerable for it. Indeed, responsibility as answerability is the notion that structures her robust

*Rational Relations View:*

*Rational-Relation Thesis:*  $\phi$  is the result of or is sustained by an individual, S's, agential activity if and only if  $\phi$  is the result of, or is sustained by, something that is itself judgment-dependent.

*Judgment-Dependence:*  $\phi$  is judgment-dependent for S if and only if S could (in principle) accept, alter, or reject  $\phi$  through reason.

The insightfulness of Smith's appeal notwithstanding, I argue that we should capture depth by rejecting the idea that being responsible is being answerable and appeal directly to Agential Self-Revelation which is already at the heart of Smith's Rational Relations View (§1). Doing so allows self-revelation accounts to accommodate the depth of moral appraisal and the fact that our responsibility outstrips our answerability.

Still, depth and a broader focus than answerability are not the only dimensions of moral appraisal that beset the self-revelation picture with troubles. The picture also faces the concern that it problematically *flattens* moral appraisal. The worry is that, since the

focus of appraisal is *always* on that which reveals who the individual is as a practical agent, self-revelation views flatten moral appraisals in the following ways:

- 1: By rendering them insensitive to certain aspects of that which is the basis for the appraisal, e.g. the extent to which the deliberation was conscious, the extent to which the reasons for and against  $\phi$ -ing were accessible, etc.;
- 2: By making them *target-invariant* such that *all* appraisals target the same thing, e.g. one's character, the revelatory act, or the decision reflected in action;
- 3: By holding that appraisals are only responses to who one is as a practical agent; and
- 4: By causing it to be the case that an appraisal that identifies an agent, S, as *bad* for  $\phi$  entails that S is *blameworthy* for  $\phi$ .

Hence, the second axis of development is leavening (§2).

Gary Watson (1996) attempts to answer the depth concern and add the requisite leavening to self-revelation accounts by distinguishing between two types of moral appraisal. Smith thinks Watson's answer threatens to return a non-unified account of being responsible according to which self-revelation is a less serious, expressive relation that obtains between an agent and her behavior.<sup>123</sup> Contrary to Smith, I embrace Watson's suggestion and employ it to develop a unified view of being responsible in two ways. First, I locate Agential Self-Revelation in Watson's original answer. Second, I appeal to the flexibility of the psychology involved in appraisals with an eye to leavening them. Thus, I preserve Watson's important idea that responsibility has an aretaic face

---

<sup>123</sup> Cf. Shoemaker (2011; forthcoming a, b) who embraces a non-unified, pluralistic account in Watsonian style. Ebels-Duggan (2013) also adopts a non-unified account, though seemingly for unrelated reasons.

and an accountability face without adopting pluralism about being responsible.

More specifically, then, on my self-revelation picture, an individual reveals their practical identity by manifesting her will. The appraisal which is involved in recognition responses or instances of holding individuals responsible comments on the will's quality. These recognition responses involve, then, recognition of an individual revealing who she is as a practical agent. And, that recognition motivates a change in/reinforcement of the appraiser's praise/blame-characteristic behavioral and attitudinal dispositions. I refer to this picture as, *the Recognition Account*. According to the Recognition Account, normative appraisals are *deep* and so no more and no less civilized when essentially tied to beliefs rather than emotions like resentment, indignation, guilt, and anger.<sup>124</sup> Further, the propriety of the recognition of will is not just determined by whether attribution is proper. Rather, the fitness of an appraisal is also sensitive to the tonal nuances brought out by concerns (1) – (4). So, by being flat about attributability, the Recognition Account maintains the *depth* and *breadth* of normative appraisals as appraisals that speak to a single conception of being responsible. I turn now to begin my sketch of the recognition account by making sense of the depth of moral appraisal.

## 1: SELF-REVELATION AND DEPTH

Suppose your teaching assistant ruins your course. She cancels discussion sections behind your back, fails to properly prepare for those she holds, and does a lackluster job of commenting on student papers. As the instructor, you took steps to address the

---

<sup>124</sup> I argue for the view that cognitive states alone can be involved in recognition responses or instances of holding responsible in Thompson (unpublished(a) & unpublished (b)).

concerns that came to your attention throughout the semester. When asked, your TA assured you that she would make the changes requested. But, to your chagrin you discover that her assurances were all disingenuous. She took no aim to do better, continued to cancel sections and office hours, and did a woefully inadequate job grading. For this you resent her. In her actions you recognize personal betrayal and disrespect toward you, the students, and the department that employed her. You feel that she shirked her duty to help you facilitate the intellectual growth of those enrolled in your course. Reflecting further you are able to identify certain loathsome character traits that you suspect are aspects of her and motivate much of her behavior. Your blame in this instance takes issue with *what she did* to you and your students. When a colleague apprised of the situation says, “It’s too bad that this happened to you, that *she* happened to you”, you think that the warm condolence does not quite capture the thought that your TA is blameworthy nor help you recover. It’s not that there is just some event the occurrence of which was bad or that there was some identifiable cause of that event that needs to be replaced. Rather, your TA recognizably manifested ill will toward you, the students, and perhaps others. To show understanding you need a colleague or someone to say something about your TA’s failure as an *agent* or her *agential act* of callous disregard for others.

Reflecting on such stories gives us a sense of Susan Wolf’s contention that when praising/blaming,

...we are doing more than identifying [an individual’s] particularly crucial role in the causal series that brings about the event in question [...] we are not merely judging the moral quality of the event with which the individual is so intimately associated; we are judging the moral quality of the individual herself in some

more focused, noninstrumental, and seemingly more serious way (Wolf, 1990, 41).

Wolf's thought nicely illustrates why we think your colleague misses the point when she simply makes a remark about the quality of the event that happened where she identifies your TA as a crucial part of the unfolding of that event. Indeed, as Wolf makes clear, the attitudes involved rest, at least in part,

...on the belief that in acting as she did she was failing to express an attitude of respect or consideration that we think she should have had [...] our resenting her rests on the belief that her acting in a way that expresses this attitude toward us (or that reveals the absence of a different attitude toward us) was up to her [...] we believe she had a kind of control over what she was doing that would have allowed her to take the attitude that we feel was our due" (Ibid., 20).

Wolf is pointing out the difference between superficial appraisals and appraisals with depth. Appraisals are *superficial* where they merely lament the unfortunate nature of the event or do so in a way that identifies the main cause of it as having a poor feature. Appraisals are *deep* where they speak to an individual's activity as an agent.

At the heart of Wolf's classic concern about depth, then, is

*Depth:* An appraisal is deep only if, at bottom, (a) the appraisal is about her activity *qua* agent; and (b) the appraisal is warranted by an individual's agential activity.

Wolf presses the concern against self-revelation views. As Wolf understands things, on self-revelation accounts, moral appraisals speak to the quality of an *aspect* of the agent rather than speaking to the quality of the agent's action as an *agential activity*. Thus, according to Wolf, self-revelation accounts do not capture the *depth* of holding responsible, since the warrant for appraisals on these accounts is not bound to whether

one possessed some kind of agential control over her decision. As the story about the TA makes salient, your colleague's attempt to show support would not have been met with dissatisfaction were blame simply a matter of noting the poor characteristics of a random occurrence. Instead, we would nod with the colleague, perhaps thinking that they were in step with our thoughts about the TA and the course. The fact that we do not see nodding along as appropriate to the situation at hand elucidates the dearth in understanding moral appraisal as merely judging the quality of some chance event.

In the remainder of this section, I develop two responses to Wolf in order to show that self-revelation views have the resources to accommodate the depth of moral appraisal. First, I develop Watson's response in light of Smith's critique. Second, I develop Smith's response in light David Shoemaker's critique. The result is an account of the depth of moral appraisals centered on agential control as a requirement for something's being self-revealing.

### 1.1: Developing Watson's Response

Watson identifies two different types of appraisal, namely, *accountability appraisals* and *aretaic appraisals*. Each type of appraisal appears to involve a distinct way of being deep. On Watson's view, accountability appraisals are deep because they are about something that is the fault of the agent who is being appraised. For him, these are the type of appraisals at the center of Wolf's critique. Alternatively, aretaic appraisals, for Watson, are deep because they are about expressions of the agent's values, her evaluative

commitments, the ends with which she identifies, etc.—i.e. they are about expressions of the agent’s moral capacity.<sup>125</sup>

Watson is sometimes read as providing a bifurcated account of being morally responsible.<sup>126</sup> The thought is that for Watson an agent’s being responsible for an action sometimes entails that the action was under the agent’s control and so is the appropriate subject of accountability appraisals. But, other times it entails that the action expresses who she is as a practical agent and is the appropriate subject of aretaic appraisals.<sup>127</sup> And, furthermore, the thought continues, where one is the fit subject of an accountability appraisal she is not always the fit subject of an aretaic appraisal and vice versa.

But, there are better readings of Watson’s idea available. To see this, consider Smith’s project and critique of Watson. Smith sets out to provide an account on which *voluntary* control of  $\phi$  is not a precondition of fitting appraisals of S for  $\phi$ . And, Smith assumes that her view and a view like Watson’s are incompatible. She writes:

Though Watson vigorously defends real self views [i.e. self-revelation views] against Wolf’s charge of “superficiality,” he nevertheless concedes that the kind of blame licensed by these accounts (which he calls “aretaic blame”) is different from (and less serious than) the kind of blame (“accountability” blame”) that he says is the focus of most traditional accounts of moral responsibility. If this is correct, then it would turn out that the sort of non-volitional account of responsibility that I [Smith] and others have [has] sought to defend does not really

---

<sup>125</sup> Watson (1996, 270 – 271).

<sup>126</sup> Smith (2008) seems to read him in this manner. Cf. Shoemaker (2011).

<sup>127</sup> Watson distinguishes between S holding an individual responsible and S making an aretaic appraisal of S. I agree the ‘holding responsible’ is somewhat unfortunate verbiage, since it seems to favor overt acts of blame. However, we should not put too much weight on this fact of language. The common thing we’re discussing are responses to manifestations of will and these responses are at bottom recognizing that will and structuring certain of parts of our dispositional profile. So, more neutral language might call them ‘recognition responses’. I thank Matt King for the discussion that led to this suggestion.

represent a genuine alternative to traditional volition-based accounts of responsibility after all. (Smith, 2008, 369)

Smith's thought that Watson's view entails that self-revelation accounts are not genuine alternatives to traditional, volition-based accounts is underwritten by the assumption that Watson doesn't and cannot accept that control of some type or other is a constraint on the fitness of appraisals. But, that assumption is false. Indeed, Watson's own words suggest that he thinks control is required even though his two types of appraisals target different aspects of the act or attitude for which the agent is being held responsible.<sup>128</sup> Watson writes:

I do not mean to say that the notion of control is irrelevant to responsibility on the self-disclosure view. Rather, issues of control are subsidiary to issues of attributability. Control bears on responsibility only so far as its absence indicates that the conduct was not attributable to the agent. "I couldn't help it" negates responsibility, for example, only by indicating that the individual's behavior (or omission) was not after all an exercise of "moral capacity". (Watson, 1996, 272, footnote omitted)

An account on which *conscious, voluntary* control of  $\phi$  is not a precondition of fitting appraisals of S for  $\phi$  could still be an account on which control *of some type or other* is a precondition. Thus, Watson and Smith can agree that the conditions for something being attributed to an individual as the basis of moral appraisal are the same despite the Janis-faced nature of moral appraisals. On such a view, self-revelation appraisals always and only speak to what was within the agent's control. And, they do so by calling attention to the fact that such things express values, commitments, and ends that belong to the agent

---

<sup>128</sup> Hence, Levy (2005) is mistaken about self-revelation views. They do not generally reject the idea that control is required for attributability.

being appraised due to the fact that “they are direct exercises of [their] moral capacity” (Ibid., 272).<sup>129</sup> It’s just that some appraisals do this while focusing on something over which the agent practiced conscious, voluntary control and other appraisals do this while focusing on something over which the agent’s practice of control was not conscious. So, despite Smith’s concerns it appears that the self-revelation picture is enriched by Watson’s thoughts rather than diminished in light of them.

But, Smith also worries that Watson’s separation of aretaic blame as different from and less serious than accountability blame is to paint the self-revelation picture as a non-genuine alternative to traditional, volition-based accounts. Smith’s concern, however, can be ameliorated by enriching the self-revelation picture in two ways. One way involves leavening moral appraisals further by thinking carefully about the array of appraisals that need to be accommodated. Another way to enrich the self-revelation picture is to sharpen the concept around which the picture should be organized. I address this latter concern first by developing Smith’s account in light of Shoemaker’s critique.

## 1.2: Developing Smith’s Response

For Smith, to be *responsible* is to be *answerable*. And, an individual is answerable for something just in case it is intimately tied to her agential activity. On Smith’s view this amounts to being judgment-dependent.<sup>130</sup> Indeed, for Smith, those things and only those things that are judgment-dependent reveal who an individual is as a practical agent.

---

<sup>129</sup> Ibid., 272 n. 31.

<sup>130</sup> Smith (2008, 382 – 387). See, Smith (2005; 2012) for a fuller description of what she calls the *rational relations view*. Cf. Scanlon (1998, 18 – 22 & 269 – 272; 2008, 193 – 198).

Thus, an agent is answerable for something if and only if that something reveals who she is as a practical agent.<sup>131</sup> Understanding this, we can see Smith's response to Wolf as unfolding in two steps. First, Smith argues that appraisals (implicitly) demand that an agent answer for some particular act or attitude. Thus, such an appraisal is appropriate only where the agent being appraised is *answerable* for that act or attitude. But, second, for Smith, being answerable for an act or attitude directly indicts an agent's activity as an agent. Therefore, Smith concludes, fitting appraisals on her self-revelation account are deep, since they are about an individual's activity as an agent and are warranted by that activity. Hence, on Smith's account being answerable for  $\phi$  grounds reactive entitlement—i.e. the warrant for having and/or expressing, say, the negative reactive attitudes (e.g. indignation, resentment, and guilt).<sup>132</sup>

Unpacking Smith's appeal to the idea that an agent is always and only answerable for that which is dependent on her judgment is important. To make the case for this view, Smith notes that "moral criticism, unlike many forms of negative assessment, seems to imply something about our activity as rational agents" (Smith, 2008, 380). The idea is that to grade someone as too little to play field hockey, too stiff for gymnastics, or too intelligent to settle for a high school diploma is different from "[accusing] someone of "selfishness"" (Ibid., 380).<sup>133</sup> To do the latter is, Smith thinks, "to make a claim about her agential activity [...; it] is to claim that she has failed, either in general or in a particular instance, to give proper weight or significance to the needs and interests of

---

<sup>131</sup> Smith (2008, 371).

<sup>132</sup> To be clear, Smith actually claims that the answerability demand "constitutes the "reactive entitlement"" (Ibid., 381). I think the best interpretation is that the answerability demand is what we are entitled to issue on the basis of one's manifesting who she is as a practical agent via her agential activity.

<sup>133</sup> Cf. Scanlon (1998, 18 – 22 & 270).

others in her attitudes and actions [...; it] can only be directed to a person with regard to things that involve her rational activity in some way” (Ibid., 380 – 381). The answerability demand that Smith claims is (implicitly) present in moral appraisal, then, “calls upon the agent to explain or justify her rational activity in some area, and to acknowledge fault if such a justification cannot be provided” (Ibid., 381).

To further elaborate, Smith holds that issuing the answerability demand (implicitly) by holding someone responsible is appropriate only where the object of the demand is something that she (in principle) can justify or answer for if called upon to do so, i.e. where the object is judgment-dependent.<sup>134</sup> The thesis given above, Judgment-Dependence, generally captures Smith’s thought and defines being judgment-dependent as that which the agent could (in principle) accept, alter, or reject by reasoning about whether to accept, alter, or reject it. In particular for Smith, something is judgment-dependent just in case it is “sensitive to our (sometimes hasty, mistaken, or incomplete) judgments about what reasons we have, and [it is] generally responsive to changes in these judgments” (Ibid., 370).<sup>135</sup>

David Shoemaker (2011) rejects Smith’s view on the grounds that a unified account of being responsible like hers does not allow for the fact that there are more things that are properly attributable to us as the basis of deep, moral appraisal than the class of things for which we are answerable. Shoemaker writes:

Aretaic appraisals are familiar. We may judge of someone, in light of some action or attitude, that he is “cowardly,” “generous,” “cruel,” or a “hopeless romantic.” What occurs here is evaluative, agential predication [...] This is not mere superficial assessment (as in the deployment of predicates like “tall,”

---

<sup>134</sup> Smith (2008, 385).

<sup>135</sup> Cf. Scanlon (1998, 18 – 22 & 272).

“skinny,” and “pretty”) for it goes to the heart of one’s actual motivations qua self-expressing agent, implicating the set of psychological elements that provide an explanation of one’s motives and attitudes in normative domains. These are the elements with which one is identified, whether or not one is conscious of, or endorses, the identification in question. (Shoemaker, 613, footnote omitted)

According to Shoemaker, then, Smith faces a dilemma. Either she must take on a pluralistic view about being responsible or she must find room for deep appraisals of those psychological elements for which we are not answerable but nevertheless seem to serve as the basis of fitting and deep moral appraisals on a unified account. Shoemaker opts for the former and I for the latter.

To make room for the deep appraisals that Shoemaker highlights, we simply need to agree with him that thinking in terms of answerability is misguided. For one, as Shoemaker points out, thinking in those terms is too closely linked to Smith’s Scanlonian Contractualism. On Scanlonian Contractualism, roughly, within the domain of *what we owe to each other*, rightness of an action is a function of whether that action is aligned with a principle no one, suitably motivated, could reasonably reject. And, following Scanlon, Shoemaker notes that the moral realm as we understand it likely extends beyond what we owe to each other.<sup>136</sup>

Furthermore, the idea that someone must be able (in principle) to answer an answerability demand where this requires offering a justification for what she did seems to be either too restrictive or unhelpfully vague. The worry on the restrictive end is that it constrains that for which we are responsible to those things where the reasons for which we did them are and/or were accessible to us. And, if it is not the case that Smith’s

---

<sup>136</sup> Scanlon (1998, 6, 171 – 177, 270 – 271, & 342 – 349) and Shoemaker (2011, 612 n. 26).

account restricts us in this way, then using the idea of answerability as the central idea around which responsibility concepts are organized is unhelpful.<sup>137</sup> Thus, in order to maintain a plausible non-pluralistic account, we must organize our understanding of responsibility around a notion that is broader than the notion of answerability.

Interestingly, there remains a central notion in Smith's view apt for organizing responsibility. That notion is

*Agential Self-Revelation:* An individual reveals who she is as a practical agent via  $\phi$  if and only if  $\phi$  is the result of, is sustained by, or lingers under the management of her agential activity.<sup>138</sup>

To further motivate this way of understanding responsibility and normative appraisals, note that Agential Self-Revelation helps answer the depth objection in a manner that is unobjectionable from the point of view of those who are likely to appeal to depth in order to reject self-revelation views. To see this and to see more clearly how my account enriches Smith's, distinguish between two types of accounts of moral responsibility<sup>139</sup>:

*Expressive Accounts:* A is an expressive account of moral responsibility just in case A focuses primarily on some characterological or expressive connection between an agent and  $\phi$  in explaining why she is morally responsible for  $\phi$ .

---

<sup>137</sup> Though I will not focus on the matter, Coleen Macnamara (2013) offers a strong rejection of the claim that moral appraisals like blame issue implicit demands like Smith's answerability demands. We can reject that portion of Smith's view without rejecting anything I adopt in formulating an alternative but similar self-revelation account.

<sup>138</sup> Cf. Smith (2008, 385 – 286).

<sup>139</sup> I owe the inspiration for describing these views as I do and the terminology 'characterological' to Manuel Vargas (2013, 135). The name 'non-expressive account' is mine. This may not be exhaustive.

*Non-Expressive Accounts:* A is a non-expressive account of moral responsibility just in case A focuses primarily on something other than a characterlogical or expressive connection between an agent and  $\phi$  in explaining why she is morally responsible for  $\phi$ .

The concern about depth is central to Wolf's influential push against Expressive Accounts, which is the proper home of self-revelation views. Smith's argument reveals that Expressive Accounts need not focus *exclusively* on some expressive connection between an agent and  $\phi$  in their full explanation of why that individual is morally responsible for  $\phi$ . In particular, proponents of Expressive Accounts can identify the connection as one grounded by that which is the primary focus of some Non-Expressive Accounts. For instance, on accounts like Wolf's, the central focus is the *rational relation* between an agent and  $\phi$  where this is a relation of agential control. Agential Self-Revelation preserves the importance of *that* connection while the primary focus of my Smithian self-revelation view, the Recognition Account, remains the expression or the revelation of practical agency.<sup>140</sup>

On my view all we need for depth is Agential Self-Revelation. In the next section, I make the agential activity referred to more concrete while handling flattening-concerns. I do so by (a) adopting Smithian monism about the preconditions for attribution and (b) showing how the tone of appraisals is adjustable in a manner that speaks to their fitness. The result is an expressive account that locates depth outside of answerability, but not at the expense of non-pluralism nor at the expense of flatness.

---

<sup>140</sup> This seems to be the wisdom we gain from my interpretation of Watson. I thank David Shoemaker for suggesting that this might be the path Watson takes.

## 2: SELF-REVELATION AND FLATNESS

Recall that flatness worries arise for the self-revelation picture due to the fact that the account of moral appraisal central to them seems to be myopically concerned with that which reveals practical agency. Again, this threatens to flatten moral appraisals

- 1: By rendering them insensitive to certain aspects of that which is the basis for the appraisal, e.g. the extent to which the deliberation was conscious, the extent to which the reasons for and against  $\phi$ -ing were accessible, etc.;
- 2: By making them *target-invariant* such that *all* appraisals target the same thing, e.g. one's character, the revelatory act, or the decision reflected in action;
- 3: By holding that appraisals are only responses to who one is as a practical agent; and
- 4: By causing it to be the case that an appraisal that identifies an agent, S, as *bad* for  $\phi$  entails that S is *blameworthy* for  $\phi$ .

Smith reads Watson as attempting to address depth concerns by appealing to different conceptions of being responsible. That reading appears mistaken. Rather, Watson seems to offer a view similar to Smith's that ties *all* attribution to a type of agential control. Watson also seems to be concerned to leaven appraisals without changing attribution conditions. For instance, Watson seems to hold that the tone of aretaic appraisals differs from the tone of accountability appraisals due to a difference in their fitting targets.<sup>141</sup> In this section, I show that my non-pluralistic Recognition Account can leaven in line with (1) – (4) by following Watson's lead. Thus, given that the Recognition Account (a)

---

<sup>141</sup> In that way, Watson seems to be responding to flattening concern (2).

explains why Wolf's argument fails, (b) offers a general method for leavening moral appraisals, and (c) offers a unified theory of being morally responsible, it steps ahead as superior to the pluralistic view that there are distinct types of moral responsibility.<sup>142</sup>

## 2.1: The Model for Flat-Leavening

My model for accommodating (1) – (4) is fairly simple. The basic structure involves keeping separate two important things: (I) the conditions for attributing something as the basis of moral appraisal, and (II) the tonal aspects of the appraisal itself. Both (I) and (II) contribute to determining whether the appraisal is fitting. But, focusing on (II) allows us to leaven so as to accommodate (1) – (4) without appealing to different conceptions of responsibility. So, the basic strategy is to tease apart (i) that which connects  $\phi$  to the agent in a manner that grounds the attributability of  $\phi$  to her as the basis of moral appraisal from (ii) other parameters to which the fitness of appraisals is sensitive.

To bring out an important complexity and make the model more concrete a further addition is needed to give shape to the following notions: resulting from, being sustained by, or lingering under the management of agential activity as it appears in Agential Self-Revelation. Consider, then, a piece of John Martin Fischer & Mark Ravizza's influential theory of control.<sup>143</sup> Fischer & Ravizza identify the following:

---

<sup>142</sup> The result might would be a view on which, roughly, that which grounds an individual's being morally responsible for  $\phi$  where  $\phi$  is one type of thing is distinct from that which grounds an individual's responsibility where  $\phi$  is a different type of thing. Cf. Shoemaker (2011, 615 – 632; forthcoming).

<sup>143</sup> To be clear, I am not committing myself to Fischer & Ravizza's view. I am simply using it for illustrative purposes.

*Guidance Control:* S has guidance control over  $\phi$  if and only if  $\phi$  is responsive to S's reasons.<sup>144</sup>

As a rough sketch,  $\phi$  is responsive to S's reasons where reasons can be the input of a process that appropriately weights them in light of the agent's web of beliefs, desires, preferences, etc. and the output of that reasons-responsive process is a verdict about  $\phi$  that is efficacious in either producing  $\phi$ , sustaining  $\phi$ , or rejecting  $\phi$ .<sup>145</sup> Thus, a lingering  $\phi$  is such that it is under the agent's guidance control lingering as it were until the agent exercises her capacity to render a verdict on  $\phi$ . Further, the exercise of guidance control can be either voluntary or non-voluntary. So, on this model, guidance control over  $\phi$  is all it takes and is necessary for  $\phi$  to be attributable to an individual as the basis of moral appraisal, since the only things that express one's practical agency and ground reactive entitlement are those elements over which the agent has guidance control.

## 2.2.: Applying the Model

Start with flattening-concern (1). According to (1), appraisals should be sensitive to how accessible the reasons for and against  $\phi$ -ing were to the agent. So, consider two individuals, Ali and Barbro. After careful deliberation where the relevant beliefs, desires, preferences, and expectations were clear to him as was their role in that deliberation, Ali decides not to allow Chinua to vote in the local election on the grounds that Chinua's race

---

<sup>144</sup> This is a rough characterization modeled on Fischer & Ravizza (1998, 31). They contrast it with:

*Regulative Control:* S has regulative control over  $\phi$  if and only if S has guidance control in the actual sequence that leads to S's doing  $\phi$  and in at least one alternative sequence that leads to S's not doing  $\phi$ —(i.e. S can do other than  $\phi$ , where doing other than  $\phi$  is having a *dual power* of guidance control with respect to  $\phi$ ).

<sup>145</sup> Fischer & Ravizza (1998, 71 – 73).

makes him filthy. Barbro, who has the same set of beliefs, desires, preferences, and expectations as Ali, also denies Chinua his right to vote. She doesn't consciously realize it, but the reasons motivating her to deny Chinua his right to vote are that his race makes him filthy. Suppose Ali and Barbro are blameworthy for denying Chinua his vote. The flattening-concern in (1) is that self-revelation views cannot deliver what we need. And, what we need is an account of moral appraisal that is sensitive to the differences between Ali and Barbro regarding their differing degrees of accessibility to the reasons entering the process that resulted in a decision to deny Chinua his right, the role those reasons played in that process, and the extent to which that process was conscious.<sup>146</sup>

To see that the Recognition Account can answer (1) note that it allows at least two things to bear on whether a moral appraisal is fitting. First, that which is being appraised must be something that is attributable as the basis of moral appraisal. Thus, it must be something over which the agent has guidance control. Second, fit appraisals have a *fitting tone*. To have a fitting tone involves a number of things. Regarding (1), a fitting tone is sensitive to things like the extent to which the agent's deliberation with respect to  $\phi$  was conscious, the extent to which the reasons for and against  $\phi$ -ing were accessible to the agent, and so on. These factors may be measured along various continuums like a *consciousness-continuum* and *accessibility-continuum* where Ali could stand as a paradigm of one extreme of each continuum while Barbro stands at the opposite extreme.<sup>147</sup> On the model, anything not under the agent's guidance control is too remote

---

<sup>146</sup> The concern is *not* that one is more infuriating than the other. Indeed, it may be that each is just as infuriating. *Neither* is the concern that we cannot distinguish a further thing for which Ali is blameworthy. Indeed, it may be that Ali is *also* blameworthy for his reflective deciding to deny Chinua the vote.

<sup>147</sup> I do not mark other continuums for the sake of simplicity. Another one might be the actual-reason continuum, which marks the extent to which an individual reflectively/consciously took as her reason(s) for

to be relevant to a deep appraisal of who she is as a practical agent. Psychological elements that lie beyond the agent's guidance control are just features and/or blemishes that remain no matter what verdict the agent returns. Such things are like her height and eye color insofar as appraisals of them are not deep and do not tell us who she is as a practical agent. Thus, the Recognition Account answers flattening-concern (1) without sacrificing depth or unification.

Answering the remaining flattening-concerns follows the pattern used to address (1). The basic idea is that appraisals have pluralistic fitness conditions. One makes them appear flat—namely, *the attributability condition*. The others elucidate their richness—namely, *the tonal conditions*. To further illustrate, consider flattening-concern (2) according to which self-revelation views are committed to appraisals being target-invariant.<sup>148</sup>

Allowing tone to be sensitive to what the *target* of the appraisal is and allowing that an appraisal's target can shift from one context to the next answers the concern. For instance, Dario's, Edwin's, and Fatima's blame may target different things. Dario's blame may target Gena's lack of courage. Edwin's blame may target his brother, Hugh's, failure to listen to and heed Edwin's expressed wish that Hugh make an effort to remember his birthday. And, Fatima's blame may target Igor's intentionally disrespectful comment. In the first case, the target seems to be an individual's *character* while Edwin's and Fatima's blame seem to target *judgments*. And, pushing further, we

---

acting the reason(s) that actually supported and motivated her decision. Or we might measure the extent to which the reasons are historically related to the agent's past deliberations and/or actions. In general, I take it that these continuums measure *agential closeness*, since they measure in some relevant sense how *close* a particular  $\phi$  is to an individual's agential activity.

<sup>148</sup> Shoemaker's (*forthcoming a*) made this worry more salient to me. I owe the insight of putting it in terms of shifting targets to him.

might think that Hugh's judgment and Igor's judgment are importantly different. Hugh's judgment might simply reveal that he is not inclined to make the relationship with his brother all that it can be. Alternatively, Igor's judgment may reveal that he is not inclined to avoid flouting moral norms.<sup>149</sup> Again, it is important to keep distinct that which renders it appropriate to *attribute* something as the basis for moral appraisal and other aspects of the appraisal that change its tone. The target-variance speaks to the fitness of tone while the fact that the target reveals who the individual is as a practical agent speaks to the fitness of its attribution as the basis of moral appraisal.

Furthermore, since the appraisal is part of the recognition of will that motivates a change in or reinforcement of the appraiser's praise/blame-characteristic attitudinal and behavioral dispositions, appropriate tone is a matter of being disposed in the right way. For example, if the target is an individual's character, then perhaps the appraiser should only be disposed to admire, have esteem for, disdain, be contemptuous toward, or be shameful of the individual who is the subject of her appraisal.<sup>150</sup>

Flattening-concern (3) is closely related to target-variance. Underwriting flattening-concern (3) is the idea that there is a difference between the *target* of a moral appraisal and *that to which the appraisal is a response*. And, the thought continues, just as appraisals are target-variant they are *response-variant* as well. The worry summarized in flattening-concern (3), then, is that on the Recognition Account moral appraisals are always *and only* responsive to an individual's agential activity.

---

<sup>149</sup> Cf. Shoemaker (2011, 615 – 632; forthcoming a, b).

<sup>150</sup> Shoemaker (forthcoming a, 104) lists targets and what he thinks are their proper emotional responsibility responses. I borrow from his list, but I take no stand on whether his list is accurate. I only agree with him that a range of responses may be fitting with respect to one target while another range of responses is fitting for another. Thus, I adopt his pluralistic view of what quality of will comes to.

I agree that there is some sense in which moral appraisals are always responsive to one's agential activity on the Recognition Account. Indeed, if S's moral appraisal of U was a response to something over which U lacked guidance control, then S's moral appraisal of U is unwarranted. However, the same point that was made above about distinguishing between that which renders  $\phi$  an appropriate basis for moral appraisal and other aspects of the appraisal can be made here. The fact that S has guidance control over  $\phi$  makes  $\phi$  attributable to U as the basis of moral appraisal. However, on the Recognition Account, S is also responding *in particular* to the fact that  $\phi$  manifests a certain quality of will. Furthermore, where one takes up a pluralistic view about what qualities of will are, we generate more response-variance.<sup>151</sup> For instance, in one context, S might be responding to the fact that  $\phi$  manifests U's poor character while in another context S is responding to the fact that  $\phi$  manifests U's disregard for an individual's interests in deciding to  $\phi$ . Thus, it is possible that, where S responds to what she takes to be a manifestation of disregard by U, S's blame of U is inappropriate even if S manifested a poor character.<sup>152</sup> Therefore, the Recognition Account has the resources to answer flattening-concern (3).

The final flattening concern, (4), is motivated by the idea that adequate views of moral responsibility must distinguish the bad from the blameworthy. For instance, Neil Levy (2005) directs our attention to the oft discussed case of Robert Harris. According to Watson's (1996) recounting of the tragic tale, Harris was actively ignored and resented

---

<sup>151</sup> See, e.g., Shoemaker (forthcoming a, b).

<sup>152</sup> Likewise, if S's response is to U's quality of will understood as U's quality of character, then that response is fitting only where S is motivated by it to have the praise-/blame-characteristic attitudinal and behavioral dispositions legitimated by an appraisal with that response-history. 'Quality of character' is borrowed from Shoemaker (forthcoming a, 100).

by his mother as a child. He spent a great deal of his teenage years in a juvenile detention center where he was repeatedly sexually and physically assaulted. During his late teenage years and early twenties, he was arrested several more times for the brutal torture of animals. Finally, a short time after being released from an adult prison facility, Harris shot two teenagers in cold blood. His brother, Daniel Harris, who was present at the shootings details how Robert laughed about the slaying of the two boys while eating the lunch they'd bought prior to being shot.

Levy contends that we might want to say that Harris is a bad person, since the attitudes expressed are in some important sense bad and genuinely his.<sup>153</sup> Still, Levy continues, we also want to allow that Harris is not blameworthy for the actions (omissions, attitudes, etc.) that reflect those bad attitudes, since he did not have the chance to become a better person.<sup>154</sup> But, Levy claims, self-revelation views do not have the resources to make that distinction. According to Levy, if the bad attitudes Harris expresses are genuinely his then self-revelation views are committed to the claim that Harris is blameworthy. Levy is pushing that on self-revelation views (a) fitting appraisals that grade an aspect of the individual that expresses who she is as a practical agent entail (b) corresponding fitting appraisals that do more than appraise an individual in light of that aspect.<sup>155</sup> In other words, Levy is claiming that self-revelation accounts are committed to the view that *aretaic appraisals* entail *non-aretaic appraisals*.

---

<sup>153</sup> Levy (2005, 5).

<sup>154</sup> Levy (2005, 5).

<sup>155</sup> Smith (387 – 390) essentially bites the bullet. For perspicuity, Smith (2008, 387 – 391) takes on an alternative challenge. On her reading of Levy, the challenge is distinguishing between the bad and the morally responsible. In response, Smith argues that very few are bad *and* not morally responsible. On her view, failure to be morally responsible entails that one is not a participant in the moral community and typically one who is bad *is* such a participant. Furthermore, Smith thinks that Levy's alternative is worse,

At least two things should be said in response. First, it is not clear that *all* aretaic appraisals are *deep* appraisals.<sup>156</sup> Thus, it is not clear that all aretaic appraisals are appraisals of an aspect over which the individual being appraised has guidance control.<sup>157</sup> Second, there is room on the Recognition Account for deep aretaic appraisals that do *not* entail deep *non-aretaic* appraisals.<sup>158</sup> It is important to keep in mind that all deep appraisals are about an aspect attributable to the agent on the grounds that it was under her guidance control, and so expressive of who she is as a practical agent. Thus, again, all deep appraisals are fitting only if they are about aspects of this sort. But, just because all appraisals share a condition necessary for being fit one moral appraisal need not entail another. For example, though an agent might be the fit subject of an appraisal with tone A, she may not be the fit subject of an appraisal with a different tone, B. Specifically, since the tone of aretaic appraisals differs from non-aretaic appraisals, it should be clear that aretaic appraisals do not entail non-aretaic appraisals. So, I might fittingly respond to my lackluster TA by thinking that she lacks integrity and hence has a character that is

---

since it requires us to view individuals as “passive victims” of their characters. So, Smith marks out a distinction between types of appraisals—*participant types* and *objective types*. In order that a *participant appraisal* is fitting, that which is being appraised must be properly attributable to the individual as a basis of moral appraisal. *Objective appraisal* requires no such thing.

<sup>156</sup> Smith makes a similar point. Smith is attempting to get clear on Watson’s distinction between aretaic appraisals and accountability appraisals. She points out that on one interpretation of aretaic appraisal such an “assessment need not involve condemnation, and therefore need not involve the idea that the person in question is at fault in some way for his or her conduct” (Smith, 2008, 377).

<sup>157</sup> Smith adds that they must be appraisals that are significant from the perspective of the normative realm toward which they are oriented. For instance, being significant from the perspective of morality requires that the content of the appraisal include the way in which S by  $\phi$ -ing has followed or flouted what is morally expected of her. And, Smith is particular here, given her adoption of a contractualist theory of normative ethics. See Smith (385 n. 21). This addition is consistent with my account of depth, but I don’t think we have to agree with Smith that her Scanlonian contract theory correctly sets out all that is or can be morally expected of her. Cf. Shoemaker (2011, 612 n. 26).

<sup>158</sup> This seems to mark a departure from Smith’s account. See Smith (2008, 385 – 390).

less than excellent. But, this does not entail that my TA is the fit subject of appraisals that, say, do more than target or respond to my TA's character.

Therefore, Levy might be correct that it is appropriate to morally appraise Harris where this is a comment about his character but inappropriate to morally appraise Harris's decision to shoot the teenagers because Harris is, say, unable to flout moral norms in the requisite manner. But, a commitment to my self-revelation view does not say otherwise. Of course, since it is true that, where morally significant and fitting, the appraisal means that Harris is morally blameworthy, we can still distinguish between being bad and being blameworthy in the way Levy suggests we must. Putting Levy's terminology with mine: *being morally bad* is a matter of being fit for a deep, moral, *aretaic* appraisal, and *being morally blameworthy* is a matter of being fit for a deep, moral, non-aretaic appraisal. But, that's just terminological. I'm happy to say that an individual is in some general sense blameworthy where they are fit for any deep, moral appraisal.

#### 4: CONCLUSION

Gary Watson offered the following in his response to Wolf:

But I think we should take real self views to have a different focus. They are prompted by a concern with agency or attributability, rather than with control and accountability. The significant relation between behavior and the "real self" is not (just) causal but executive and expressive. When thought or behavior are exercises of what Dewey calls an agent's moral capacity, they and their results are open to distinctive kinds of evaluation. These evaluations are inescapably evaluations of the agent because the conduct in question expresses the agent's

own evaluative commitments, her adoption of some ends among others. To adopt some ends among others is to declare what one stands for. (Watson, 1996, 270)

The Recognition Account agrees with most of what Watson says. The primary disagreement, if there is any, lies in the fact that self-revelation accounts can be concerned with control and accountability *as well as* the revelation of practical agency. By adopting Agential Self-Revelation and separating issues that speak to whether  $\phi$  is attributable to an individual as the basis of moral appraisal from issues that speak to the intonation of the appraisal, we capture the important idea that the ‘distinctive kinds of evaluation’ to which certain individuals are open ‘are inescapably evaluations of the agent because the conduct in question expresses the agent’s own evaluative commitments’, since in her decision ‘what she stands for’ is revealed.

However, the Recognition Account goes further. In light of Shoemaker’s critique of Smith, I argued that elements within an agent’s psychology are attributable as the basis of moral appraisal so long as she has agential control over them. Thus, whether she has taken a stand on them or not they speak to who she is as a practical agent and so reveal this identity when brought to light.

The Recognition Account is also general. Deep appraisal in any normative domain could be understood as appraisal warranted by one’s activity *qua* agent even if the significance of moral appraisal is different than that “of our activity in some other normative domain (such as philosophy, or music)” (Smith, 2008, 386 – 387).<sup>159</sup> Thus, this picture makes sense of deep appraisals in all normative realms important to our interpersonal flourishing as agents.

---

<sup>159</sup> I would add epistemic pursuits to this list as well.

Finally, the picture developed here also shows that the recognition account does not civilize blame in the sense that it only warrants superficial appraisals. However, looking beyond concerns about the civilization of blame, we might see trouble looming. The trouble is connected to the fact that the recognition account is committed the idea that when we hold responsible we are recognizing a manifestation of will. Thus, the recognition account is a self-revelation account that is also a *quality of will view*. I develop and defend the recognition account along that axis in the next chapter.

## VI: NEGLIGENCE AND RESENTMENT

Common wisdom has it that a negligent individual who injures another is morally responsible and blameworthy for doing so.<sup>160</sup> Thus, we tend to think that unreasonable carelessness, thoughtlessness, or inattentiveness undermines neither an individual's responsibility nor her blameworthiness. Nevertheless, many hold that negligent individuals fail to meet certain mental conditions at the moment of injury which are required for being morally responsible and, hence, for being blameworthy.<sup>161</sup> This conundrum is especially challenging for the family of accounts generally referred to as *Quality of Will Views* on which an individual is morally responsible for  $\phi$  just in case  $\phi$  expresses who she is as a practical agent by manifesting her will. Here, roughly, an agent is blameworthy just in case  $\phi$  also manifests *ill will*.<sup>162</sup> A central worry is that where an agent's negligence cannot be tied to a prior culpable act or attitude we must rely solely on this manifestation to ground her moral responsibility and blameworthiness for the injury. But, the concern continues, the most plausible interpretations of the manifestation relation forces us to assume implausibly that *all* negligent injurers harbor some morally bad attitude about the person injured or persons in general which caused the negligent behavior. In this paper, I develop a quality of will account of moral responsibility against the background of this concern.

---

<sup>160</sup> For a legal description of negligence, see, e.g. *Model Penal Code*, Sections 2.02(2)(d). Cf. Lunney & Oliphant (2001, 129). Compare, e.g., Hart (1968), Feinberg (1969), Glover (1970, 60 – 61), H. Smith (1983; 2011), Sverdlik (1993), Zimmerman (1997), and King (2009).

<sup>161</sup> Many adopt H. Smith's (1983) *tracing strategy* to account for culpable but negligent injury. See below.

<sup>162</sup> See, e.g., Strawson (1962), Scanlon (1988; 1998, chpt. 6; 2008, chpt. 4), Wallace (1994, esp. 127 – 136), McKenna (2012), Greenspan (2001), Arpaly (2003, esp. chpt. 3), and Hieronymi (2004; 2011).

One of my main contributions to the quality of will picture is an illustration of how *absences* of things like proper care or regard as opposed to the *presence* of a morally questionable motive or attitude can cause an individual to manifest a will with ill quality (§1). This paper also advances the view that an individual who lacks proper care or regard is (fully) excused or exempt *only if* that lack is not under her agential control (§2). In particular, I reject the view currently gaining popularity that one who lacks the capacity to understand praise or blame, which in some instances may be tied to lack of proper care or regard, fully immunizes them from our responsibility responses.<sup>163</sup>

I support the latter contribution by appealing to the nature of the psychological attitudes involved in our responsibility responses. Throughout I lean on my empirically informed Recognition Account of responsibility responses according to which, an individual S responds to an individual U as responsible just in case S recognizes or judges that U manifested good or ill will.<sup>164</sup> Furthermore, though sometimes non-affective, that recognition of will motivates a change in or reinforcement of S's praise/blame-characteristic, behavioral and attitudinal dispositions.<sup>165</sup> The idea in a nutshell, then, is that these recognition responses are fitting even if they are responses to individuals who cannot understand the interpersonal significance of them. Only those who cannot manifest a will with certain morally significant qualities are (fully) excused or exempt, i.e. not fit subjects of such responses. Before advancing that view, I begin by accommodating the common wisdom that negligent injurers are morally blameworthy.

---

<sup>163</sup> For early proponents, see Oshana (1997; 2004) and McKenna (2012). For an example of recent converts see Pereboom (2013; 2014) and compare Shoemaker (2011).

<sup>164</sup> [omitted].

<sup>165</sup> My account is inspired by but different from P. F. Strawson's (1962). [omitted]

## 1: NEGLIGENCE AND QUALITY OF WILL

Traditionally, cases involving negligent injurers are handled by locating some non-negligent act performed by them that brought about their negligence.<sup>166</sup> Call these cases, *tracing cases*, since we account for the common wisdom that the agent is responsible by tracing the origin of their current negligence to a prior event. However, there are also cases that some take to involve negligent injurers where, supposedly, none of their prior acts brought about their lack of foresight.<sup>167</sup> Call these cases, *non-tracing cases*. For example, the following are taken by some to be non-tracing cases:

*Driveway Disaster*: Nate, tired from waking up early, backs out of his driveway. His thoughts turn to his meetings that day, and his attention is partly focused on a radio commercial. Due to his inattention, Nate doesn't see a child walking to school and so hits him, breaking the child's leg. (King, 2009, 578)<sup>168</sup>

*Hot Dog*: Alessandra, a soccer mom, has gone to pick up her children at their elementary school. As usual, Alessandra is accompanied by the family's border collie, Bathsheba, who rides in the back of the van. Although it is very hot, the pick-up has never taken long, so Alessandra leaves Sheba in the van while she goes to gather her children. This time, however, Alessandra is greeted by a tangled tale of misbehavior, ill-considered punishment, and administrative bungling which requires several hours of indignant sorting out. During that time, Sheba languishes, forgotten, in the locked car. When Alessandra and her children finally make it to the parking lot, they find Sheba dead from heat prostration. (H. Smith, 2011, 2, footnote omitted)<sup>169</sup>.

---

<sup>166</sup> H. Smith (1983) dubs these prior acts 'benighting acts'. See Vargas (2005, 269 – 271) for a nice summary of this tradition.

<sup>167</sup> See, e.g., Sverdlik (1993, esp. 137 & 143 – 145), Vargas (2005, 271 – 273) and Sher (2009, chpt. 2). Sverdlik (1993, 141) calls this 'pure negligence'.

<sup>168</sup> H. Smith (2011, 117) also offers Driveway Disaster.

<sup>169</sup> Hot Dog is a variant on a case presented by Sher (2009, 24).

Driveway Disaster and Hot Dog are non-tracing cases involving a negligent injurer.<sup>170</sup> Thus, we're assuming that the actions of Nate and Alessandra are acts "in which the agent fails to advert to some risk of harm and so acts in a way that produces or could produce this harm" (H. Smith, 2011,116), so "their conduct was unreasonably risky" (King, 2009, 578).<sup>171</sup> And, we'll assume with the common wisdom that, at least intuitively, Nate and Alessandra are morally responsible and blameworthy.

Non-tracing cases put the heart of quality of will theories to the test, *viz.* the manifests relation. The primary significance of failing to account for the moral responsibility or blameworthiness of individuals like Nate and Alessandra by relying on the idea that they manifest a particular quality of will at the moment of injury is that this failure would raise doubts about relying on the idea elsewhere. So, to assess whether the manifests relation withstands this test I will consider a particularly crisp articulation of an argument meant to show that quality of will views fail in this regard offered by Matt King (2009).<sup>172</sup> Prior to considering that argument I want to characterize the way in which quality of will views should generally handle cases involving those who negligently injure.

---

<sup>170</sup> H. Smith (2011, 117) wonders whether there actually are any non-tracing cases of negligent injury. Whether they exist or not is not my concern, since I argue for a way to account for them if they exist.

<sup>171</sup> It is debatable which features, if any, distinguish negligence cases from recklessness cases. According to *Model Penal Code* 2.02(2)(c) and 2.02(2)(d), the former involves a certain lack of awareness of the likely consequences whereas the latter requires awareness of such risks. For discussion, see e.g. Terry (1915), Edgerton (1926), White (1961), Brady (1980; 1996), Baker (1987), and Husak (2011). For an interesting global comparison, see e.g. Heller & Dubber (2011).

<sup>172</sup> H. Smith (2011) runs King's argument against a certain quality of will view.

## 1.1: The Basic Idea

The basic idea is fairly straightforward. On quality of will views, one can act in a manner for which she is blameworthy where the act manifests her will and the will is morally objectionable not because of something in its content but because of something its content *lacks*. In cases of negligent injury, the will is of ill quality insofar as it shows that one was *not* being reasonably circumspect about the activity in which she was engaged. And, this failure of reasonable circumspection can serve as the rational grounds for holding the agent responsible. To see this, consider the following case adapted from David Shoemaker (2011, 620):

*Chores:* George and Martha have been married for several years. Throughout the year, Martha has to drop subtle but increasingly forceful hints that George is failing to do his share of the house work. These should merely be hints, she thinks, because were she to have to tell George explicitly to help and why he should, his help would be less a reflection of his seeing on his own that things are unfairly disadvantaging Martha than merely another task that, as husband, he must perform. Unfortunately, each year George misses the hint. Despite the fact that there are plenty of household chores to be done, George spends his time watching the game, reading, working on a puzzle and other leisurely activities. Of course, he'll clean the dishes after breakfast or pick up a stray sock on his way to his office. Martha feigns gratitude. She is actually mildly resentful. She wonders why George is so unreasonably careless, thoughtless, or inattentive when it comes to the fair division of domestic labor, given that he is otherwise very attentive to his familial obligations. Recognizing in his act a lack of due diligence she is no longer disposed to be as chipper about their weekends, nor is she disposed to pay as much attention to George's hints about what he would like to do with her in their free time.

In *Chores*, George's doing little to help balance the amount of household work done

appears to be negligent and Martha's reaction seems natural and fitting. But, the will manifested by George's act carries no objectionable regard toward Martha in its content nor does it issue from a lack of regard for Martha in particular or humanity in general. Rather, the will manifested seems to simply *lack* something in its content—namely, it fails to say that George was reasonably circumspect about the amount of domestic work he should do. And, as Shoemaker points out, this sort of failure makes their relationship “worse than it could have been” (Shoemaker, 2011, 622). Thus, Martha's mild reaction whereby she holds George responsible is appropriate to the occasion.

In general, quality of will theorists should treat genuine cases of negligent injury in the same way. Individuals like Nate and Allesandra seem to act in ways that establish rational grounds for holding each responsible just as George's act seems to introduce reasons for Martha to hold him responsible. So, the basic idea is that typically one's negligent injury manifests a will with ill quality due to the fact that it fails to meet a certain standard. This is no different than non-negligent injury where an individual purposefully aims to flout a moral tenet. But, in the case of negligent injury the quality fails to meet the standard in virtue of the fact that it says that the actor was not undertaking the activity in a particular way where failure to do so flouts a moral imperative. Each grounds reason to hold the actor responsible, though perhaps being unreasonably attentive is grounds for a more mild form of blame than being purposefully immoral. But the question remains: Can the manifests relation do the work to which the quality of will theorists puts it in cases of negligent injury?

## 1.2: The Adequacy of the *Manifests Relation*

King argues that all theories of moral responsibility fail to account for the moral responsibility of individuals like Nate and Alessandra. Thus, he attempts to substantiate

*The Rejection Thesis:* We ought to reject the claim that negligent agents are responsible for the harms they bring about.<sup>173</sup>

King's argument in support of the rejection thesis begins with

*The Rejection Dilemma:* If we are unable to give a general explanation of responsibility that can account for cases involving negligence, either negligence requires an exceptional *ad hoc* explanation, or the rejection thesis is true.<sup>174</sup>

He follows by arguing that the antecedent is true and that we are theoretically better off if we accept The Rejection Thesis. Along his way to arguing that no one can offer a general explanation of responsibility that accounts for negligent individuals like Nate and Alessandra, King rejects the explanation he takes to be offered by quality of will theorists.<sup>175</sup>

King misinterprets the theories he sets out to impugn. To show this, the discussion here elucidates a more perspicuous manifests relation and thereby further fills out quality of will views in two ways. On the one hand, it illuminates the basic idea delineated above regarding the presence of absence. On the other hand, it puts us in a

---

<sup>173</sup> King (2009, 582)

<sup>174</sup> Everything before the 'or' is King (2009, 582).

<sup>175</sup> H. Smith (2011) endorses King's argument against a particular subset of quality of will views—namely, those that follow A. M. Smith's (2005; 2008) construal of such a view.

position to address whether quality of will theorists should agree that negligent injurers are blameworthy. In the next section, I argue that they should so long as they have the capacity to manifest ill will.

King characterizes quality of will views as those views on which the fact that “an agent’s conduct manifests an ill quality of will is sufficient for demonstrating responsibility” (King, 2009, 583). He explains how these views might be able to accept that negligent injurers like Nate are morally responsible:

It doesn’t matter that Negligent Nate doesn’t consciously choose to risk running the child over, his failure to pay adequate attention nevertheless manifests a quality of ill will. Naturally, it’s not as bad a quality of will as if he had harmed the child intentionally or knowingly, say. Nevertheless, failure to pay attention when one is engaged in activities that pose a risk of serious harm displays a lack of consideration for those who you risk harm to. So, Negligent Nate displays ill will toward the child, and is thus responsible for harming him. (King, 2009, 583)

King’s critique of the quality of will position is basically a call for an investigation into the “manifests relation”. He follows by offering two interpretations of the relation neither of which, King argues, do the work necessary to account for the responsibility of folks that injure negligently.

The interpretations of the manifests relation King suggests are

*Evidential Reading:*  $\phi$  manifests ill will just in case  $\phi$  is evidence that S’s will is ill.

*Causal Reading:*  $\phi$  manifests ill will just in case  $\phi$  is caused by an ill quality of S’s will.

Consider, first, King’s argument against the Evidential Reading interpretation. Call this argument *The Evidential Argument*. According to King, only a *pattern* of behavior

would serve as sufficient evidence that an individual possesses an ill will.<sup>176</sup> King concludes that quality of will views that take up the Evidential Reading interpretation of the manifests relation can at best only account for agents who have a particular pattern of behavior. At worst, King presses, they cannot account for the intuition that an individual is blameworthy for any *single* negligent act.<sup>177</sup> Thus, King draws the conclusion that the Evidential Relation interpretation is unhelpful.

King's Evidential Argument is unsound. First, King's argument seems to suggest that 'will' as understood by the quality of will theorist refers only to a characterological aspect of the agent. But, contrary to King's suggestion, quality of will theorists are not univocal about what 'will' refers to. For instance, 'will' can refer to a characterological aspect of the agent and it can refer to a particular decision that the agent made.<sup>178</sup> Second, King assumes that the quality of this presumably characterological aspect of the agent can only be properly assessed by looking at a number of acts that manifested the agent's will or that aspect. But, even if the will referred to in quality of will accounts was some characterological aspect, that assumption is false. If Jaclyn walks out of her house in the suburbs and starts pointing and shooting her handgun in random directions, we have good evidence that Jaclyn's will is ill.<sup>179</sup> In this case, that her behavior was part of a pattern would simply be more evidence of the quality of her will. Furthermore, it seems that we could say of any single act  $\phi$ , if  $\phi$  is part of the series of acts correctly taken as evidence that S's will has a particular quality, then that  $\phi$  is also evidence that such is the case and thereby manifests it. Where someone else, U, sees S  $\phi$ , whether U is aware that

---

<sup>176</sup> King (2009, 584).

<sup>177</sup> Ibid., 584.

<sup>178</sup> [omitted] Cf. Shoemaker (forthcoming a, b).

<sup>179</sup> I thank [omitted].

$\phi$  makes up a pattern of behavior that would itself be evidence that S's will is ill,  $\phi$  is evidence for U that S's will is ill.

Perhaps, however, two things speak in favor of the view that the Evidential Reading interpretation cannot be the *whole* story even if it is part of the story. For one, the natural reading of 'manifests ill will' seems to favor a causal reading of the relation rather than an evidential one. For another, even if something like the Evidential Reading is part of the story, it seems that we need to know why it makes sense to take certain acts to be evidence that S possesses a particular quality of will and causal stories often provide this link. So, here I set aside the Evidential Reading interpretation and focus on the argument King offers against the Causal Reading interpretation.

King's objection to the Causal Reading interpretation is that it requires the negligent injurer to possess "some ill quality of will, either toward the [one injured] or in general" (King, 2009, 584). Against this requirement, King writes:

...Nate doesn't even know the child is there, or that anyone is there for that matter. Nor is he even consciously thinking about backing up, so we might seriously doubt that his current frame of mind is such that his attitudes toward others, whatever they are, would be engaged to cause his lack of care. More importantly, however, Nate could very well hold quite positive attitudes toward the child he injures. [...] If this were the case, it would be hard to claim that Nate holds any ill will toward the child, much less that that ill will caused his failure to pay attention. (Ibid., 585, footnote omitted)

So, according to King, (1) it is implausible to suppose that every negligent injurer possesses an ill quality of will like lack of concern for the well-being of others, but (2) the quality of will theorist under the Causal Reading interpretation *must* suppose that such is the case in order to account for the blameworthiness of negligent injurers.

King's argument is insightful. It does seem worrisome to suppose that *all* those who injure others negligently or otherwise must be acting in a manner that fails to bear appropriate concern for the well-being of others.<sup>180</sup> The insightfulness of this point notwithstanding, like the Evidential Argument before it, this sort of argument fails due to the fact that it rests on a false assumption about quality of will views.

The type of argument King issues leans heavily on the assumption that quality of will views must accept that the injurious act that was undertaken with an unreasonable lack of circumspection must be *caused by* an attitude of lack of concern for the well-being of either the injured or others in general. But, quality of will theorists need not accept that claim. Rather, the causally relevant aspect (i.e. the individual's agential faculties, broadly construed) and the quality of the will it produces are importantly distinct. One's agential activity produces and makes available for evaluation one's will and thus is the relevant cause. The quality of that will is another matter entirely.<sup>181</sup> Furthermore, that quality is determined by the gaps in her deliberative process just as much as it is determined by the reasons that entered the deliberative process or the attitudes brought to bear on the question of what she should do. Or, more guardedly, the gaps that are under her agential control in part determine whether her will was sufficiently good or sufficiently not good to justify a particular evaluation of it. These gaps are such that were the agent to try and be reasonably circumspect about the risks involved or the significance of acting in a particular way, then she would have acted non-negligently.<sup>182</sup>

---

<sup>180</sup> Cf. H. Smith (2011, 119 – 120).

<sup>181</sup> Cf. Paul C. Taylor (2004, 33 – 35), Arpaly (2006, 14 – 15), and McKenna (2012, 58 – 59).

<sup>182</sup> See [omitted] for more details about being under one's agential control. Cf. A. Smith (2008).

Thus, for instance, reasons entered Nate's processes that tipped the balance of reasons in favor of operating a vehicle while paying no attention to who is in its path or silenced worries that might have lead Alessandra to halt her conversation so that she could check that her dog wasn't suffocating. Thus, Nate's and Alessandra's agency produced and manifested a will. That will is, then, attributable as the basis of moral appraisal. And, the appraisal of that will is based on its quality. So, e.g., the appraisal might focus on the level of respect for particular individuals ensconced in that will. And, that level of respect is not just determined by specific attitudes of the agent. It is also set by relevant gaps in the attitudes and reasons brought to bear on what she should do. Where that respect, say, falls below a certain threshold it is considered to be of ill quality or not sufficiently good and the agent is thereby blameworthy for it.<sup>183</sup>

So, more perspicuously, perhaps the Causal Reading interpretation of the manifests relation regarding blameworthiness should be something like

Causal Reading\*:  $\phi$  manifests *ill* will just in case  $\phi$  is caused by S's agential activity which reflects a decision that is of ill quality.<sup>184</sup>

Elaborating a bit, consider how quality of will theorist, Michael McKenna analyzes a case like the following:

Suppose, for example, Casper intentionally cancels his weekend business plans, and he does so with the intention of arranging a golf outing with his buddies. The reasons for his making these plans include a desire to spend some time in the sun, talk with a few old friends, enjoy a challenging game, and forget about his

---

<sup>183</sup> Cf. Nomy Arpaly (2003, 176 – 177)—a low-quality decision to  $\phi$  does not need a corresponding low regard for it to be such and thereby manifest a low-quality of will.

<sup>184</sup> I should stress that this is just an example. Some may balk at my use of 'decision' in Causal Reading\* because they think that  $\phi$  can reflect something else that is of low quality and count as a manifestation of ill will. That is fine so far as my argument goes. The point remains that the quality is distinct from the cause.

burdens at the office. His intentions and the reasons figuring in them are innocuous. But suppose that in making these arrangements, he did not consider taking time to spend with his daughter, who has recently become quite ill. In conducting himself this way, Casper acts from a morally criticizable will, but it is not because, for instance, he intended, decided, or chose to neglect his daughter. He did nothing of the sort; she did not figure in his planning at all. (Michael McKenna, 2012, 59 – 60).

In this case, as McKenna explains, Casper manifests a morally worrisome way of attending to the reasons he has, i.e., he makes a ‘low-quality decision’. But, a low-quality decision to  $\phi$  does not need a corresponding, morally low belief, desire, preference, expectation, etc. for it to be a low-quality decision.<sup>185</sup> And, that is precisely what King’s argument assumes. Thus, arguments like the causal argument fail to impugn the Causal Reading interpretation of the manifests relation. Therefore, quality of will views can accommodate the blameworthiness of negligent injurers.<sup>186</sup>

Proponents of the quality of will picture can also accommodate the *moral responsibility* of negligent injurers in much the same way. The idea is that one manifests her practical agency wherever her behavior reflects a decision to  $\phi$  produced by her practical agency. The connection that the manifestation bears to her agency is why we correctly take such things to manifest who she is as a practical agent and, thereby, manifest a will appraisable on the basis of some quality or another of it. Thus, so long as the negligent injurer’s decision to  $\phi$  which in turn caused the injury is connected in the

---

<sup>185</sup> Cf. Arpaly (2003, 176 – 177). For instance, if we take up the view that manifestations of ill will show that one fails to show proper regard, regard is a concept too wide to be pinned down to a particular belief, desire, etc. See, e.g., Taylor (2004, 33 – 35), Arpaly (2006, 14 – 15), and McKenna (2012, 58 – 59). Also, here I elaborate what may seem like a monistic view about quality of will. However, this is all basically translatable into Shoemaker’s (*forthcoming* a, b) pluralistic account.

<sup>186</sup> It should be clear that the argument I offer stands even if we change the Causal Reading interpretation by replacing ‘is caused by’ with something like ‘reflects’.

right way to her agency she manifests her will and so is morally responsible for the injury.<sup>187</sup> So, as far as responsibility for  $\phi$  goes, we need an interpretation of the manifests relation that elucidates a relation between an item and will *simpliciter*. Here is a reasonable interpretation of the manifests relation to pair with Causal Reading\*:

*Causal Reading\*\**:  $\phi$  manifests S's will just in case  $\phi$  is caused by S's agential activity.

Given these interpretations, in order to support The Rejection Thesis King must show either (a) that a negligent injurer's actions are not caused by their agential activity, or (b) that a negligent injurer's actions reflect a will of no quality. Nothing King says supports (a) or (b). Thus, King's argument fails.

### 1.3: Further Elaboration

Despite appearances, some might think that matters have not improved. The worry might be that the fix on offer simply takes the judgment of blameworthiness as one about the lack of concern itself and not one about the injury.<sup>188</sup> King offers a version of this concern that ends in the claim that it is the failure to “consider the risk of harm [Nate is] posing to others as he backs out [...] that constitutes his negative quality of will” (Ibid., 586).<sup>189</sup>

I'm not convinced that King is correct that this sort of move “solves the matter by fiat” in that it “simply lowers the bar for what we need in order to draw the inference”

---

<sup>187</sup> [omitted]. Cf. A. Smith (2008).

<sup>188</sup> King (2009, 586).

<sup>189</sup> Perhaps this is how Sverdlick (1993, 142 – 143) would reply.

(King, 2009, 586) that Nate is blameworthy. And, I have failed “to give us a general explanation of responsibility across cases” (Ibid., 586). Regarding the former, my solution says nothing about what the judgment of blameworthiness is about. Rather, I simply point out that quality of will theorists are not committed to the view that one’s unreasonable carelessness (i.e. negligence) leading to injury of  $x$  in the cases must have been caused by  $S$ ’s harboring a lack of concern for  $x$  or  $x$ -type things in general. This response doesn’t lower the bar for blameworthiness, since it doesn’t say what it takes for negligent injurers to be blameworthy. And, regarding the latter concern King raises, I haven’t changed the general explanation of responsibility across cases at all. One is still responsible, say, for an action just in case it expresses who she is as a practical agent—i.e. just in case it manifests her will. And, where that will is ill or not sufficiently good, she’s blameworthy. Once we have a viable understanding of the manifests relation we’ve solved the sort of concern King raises. It remains an open question what it is that our judgments of blameworthiness are about.

However, perhaps King would push us to consider the argument that completes his case for excusing negligent injurers<sup>190</sup>:

*King’s Argument from Analogy:*

- K1) If negligence and inadvertence are indistinguishable from the perspective of responsibility, then either (a) negligent agents aren’t responsible/blameworthy for the harms they cause, or (b) inadvertent agents are responsible/blameworthy for the harms they cause.
- K2) Negligence and inadvertence are indistinguishable from the perspective of responsibility.

---

<sup>190</sup> King (2009, 587 – 595). (K1) is basically a direct quote from Ibid., 590.

KC1) Thus, either (a) negligent agents aren't responsible/blameworthy for the harms they cause, or (b) inadvertent agents are responsible/blameworthy for the harms they cause.

K3) It is not the case that inadvertent agents are responsible/blameworthy for the harms they cause.

KC2) Therefore, negligent agents aren't responsible/blameworthy for the harms they cause.

King supports (K3) by arguing that the blameworthiness of inadvertent injurers as with negligent injurers will involve offering some *ad hoc* explanation, "since both are characterized by the absence of a conscious element" (*Ibid.*, 591). But, I showed how the quality of will theorist can account for the moral responsibility and blameworthiness of those who injure negligently. Furthermore, the account is not *ad hoc*. Rather, it fits seamlessly into the general quality of will story. Thus, quality of will theorists can accept and explain the moral responsibility/blameworthiness of those who injure negligently. That is, they can reject (K3) and thereby occlude the move to (KC2).

Still, one might accept (K3) on the grounds that they agree with King that those who injure inadvertently are intuitively neither morally responsible nor blameworthy for the injury. King offers the following case as a case of inadvertent injury (*Ibid.*, 588):

*Leadfoot Lenny*: Lenny is at a party where a group of friends are gathered watching a movie. There are more people than seats, and some have gotten comfortable lying on the floor. Lenny gets up to get a soda from the fridge, and in the course of stepping around and over people, he inadvertently steps on his friend's hand. He didn't mean to step on his friend's hand, but he was distracted by the movie, and so he did.

King explains that he thinks “it true that inadvertence seems to count as a consideration that undermines responsibility precisely because it involves the lack of a conscious mental element” (Ibid. 590). He goes on to claim that when “one does something only inadvertently, it is an unintentional result, one the individual didn’t mean to bring about” (Ibid., 590). Accepting these explanations, then, one might press that (K3) stands as does the move to (KC2).

The main problem with supporting (K3) by relying on the idea that inadvertent injury and negligent injury involves the lack of a particular conscious element, *viz.*, *intentionally aiming to harm someone*, is that sometimes the presence of such an absence is precisely what makes the will manifested of ill quality. Nate’s, Allesandra’s, and George’s lack of due diligence says something important about the decisions reflected in their actions and grounds the appraisal that each is of poor quality.

Still, (K3) only partly rests on the argument I’ve already shown fails. King also presses that the cases seem out of step with each other from the perspective of praiseworthiness.<sup>191</sup> According to King, intuitively, inadvertent agents are neither responsible nor praiseworthy.<sup>192</sup> King writes:

An agent who brings about some good outcome inadvertently surely isn’t praiseworthy for it, no matter how great the good achieved. And even if there are important differences between praiseworthiness and blameworthiness, it doesn’t seem as if an agent who brings about some good inadvertently is even responsible for it, claims of praise aside. So while insisting that both negligent and inadvertent agents are responsible is an available option, I think such a claim both conflicts with a natural view about inadvertence and gets us no closer to explaining responsibility in either case. (*Ibid.*)

---

<sup>191</sup> *Ibid.*, 591.

<sup>192</sup> *Ibid.*

In light of this, one might argue that due to the fact that inadvertent individuals are not morally responsible (or praiseworthy) in the positive case where they bring about some good, we should deny that they are in the negative case as well—i.e., we should adopt (K3).

In response, at least three things should be said. First, King offers us no reason to accept his claims about the positive case. Second, and in relation to the first, if his claims rest on the idea that the individual in the positive cases of inadvertence are not responsible (or praiseworthy) for the good they bring about because they did not bring about that good intentionally or on the basis of a morally good or exceptional attitude, then everything I've said regarding the negative case will apply here as well. The bottom line is that the absence of a good/bad attitude or disposition by those we help/injure cannot alone tell against one's responsibility or her praise/blame-worthiness.

Finally, we could let (K3) stand and still reject (KC2) by calling (K2) into question. Notice, it is not obvious that negligent acts and inadvertent acts are on a par from the perspective of moral responsibility. For one, it might seem odd to speak of one negligently aiding someone and so being worthy of praise for it. We might wonder for instance whether acting unreasonably careless can be met with positive, normative appraisal. Likewise, it isn't clear that it is always inappropriate to feel gratitude toward someone who inadvertently knocks the 'end-of-order' bar behind their groceries on the conveyor belt thereby allowing one to unload their basket. Given the right kind of absent mindedness, that gratitude may be fitting. Thus, if it makes sense to react positively to someone who helps you inadvertently but it makes no sense to do so with respect to negligence, then (K2) is vulnerable as well.

It remains to be seen whether inadvertent injurers are indeed responsible and/or blameworthy. But, if I'm right, quality of will views can accommodate the common wisdom that negligent injurers are and by the same lights such views can account for the notion that inadvertent injurers are as well. Moreover, those who adopt a quality of will view in their attempt to develop a theory of moral responsibility that is most clearly aligned with P. F. Strawson's (1962) quality of will project are unlikely to excuse negligent injurers from blame merely on the basis of their lack of foresight. In the next section, I elucidate this thought and nod acceptance to the idea of a fairly broad community of morally responsible agents as well as a fairly broad range of actions that can manifest one's will thereby making such actions attributable as the basis of moral appraisal.

## 2: MANIFESTING WILL, NEGLIGENCE, AND THE MORAL COMMUNITY

On quality of will views, the quality of an individual's manifested will is determined by what is revealed about that person's practical agency. And, I have been arguing that manifestations of will can reflect views the agent has as well as absences and gaps in her attitudes and reasoning. Where those views and gaps are morally significant they may enter into our evaluation of the agent's will. This suggests of course that individuals who lack the capacity to accept or reject claims on the basis of reasons are outside the community of morally responsible agents. Likewise, those who cannot fill absences or close gaps on the basis of reasons are intuitively in that same segregated group—namely, the group of individuals unfit for moral appraisal. But, it also suggests that those who

practice their agential capacity to respond to reasons and thereby inexcusably generate an ill will give us *pro tanto* reason to blame them. Some quality of will theorists reject that view. I close by developing and defending it.

To guide the discussion, let's just work with a simple version of the view that we have *pro tanto* reason to blame an individual who manifests an ill will without excuse via her agential capacity to respond to reasons. On the picture defended in this paper so far, then, an individual can do this negligently by being insufficiently sensitive to the reasons she has. The simple version of this view is committed to two main ideas. First, it is necessary for an individual S to have the capacity to directly affect x through reason in order for x to be properly attributed to her as the basis of moral appraisal. Second, it is necessary for an individual S to have the capacity to directly affect y through reason in order for y to be a psychological element that figures into determining the quality of S's manifested will. Hence, absence of care or due regard are psychological elements that can figure into determining the quality of S's will only where those attitudes and what it takes to fill those absences are within the proximal reach of S's rational capacities. For instance, Nate, Allesandra, and George seem to have that control.

The theorists who reject all versions of that simple view do not stop there. They allow, say, that Nate is neither morally responsible nor blameworthy if he fails to have the capacity to understand what is going on when someone responds to individuals as a responsible agents by praising or blaming them. And, that is a lack that may inhibit due care or regard. The following case helps to illustrate their view and how it conflicts with the one I want to defend<sup>193</sup>:

---

<sup>193</sup> This is a modification of the case offered by McKenna (2012, 108 – 109).

*Crusoe*: Suppose Crusoe can apply moral predicates, and so can understand that his beating Rover is cruel. Perhaps he thinks it is morally bad and believes that what he is doing is morally wrong. Nevertheless, Crusoe has no concept of an interpersonal moral responsibility exchange. He has no idea what it would mean to be blamed, to be held to account for his actions, for him to offer an apology to others for his conduct, or for him to “take responsibility” and accept and experience guilt. That sort of conversational dynamic is simply beyond him.

McKenna, explains that blaming Crusoe “is pointless if he cannot acknowledge it as blame” (McKenna, 2012, 109).<sup>194</sup> Expanding, McKenna draws our attention to an individual named Robinson who possesses the capacities Crusoe lacks that render the latter neither praiseworthy nor blameworthy. McKenna tells us that there “is a kind of conversation others can have with [Robinson], one to which he can be expected to bear further burdens, perhaps in the form of an apology to relevant others in the moral community, an acceptance of guilt, or even a willingness to seek forgiveness in the process of moral improvement” (McKenna, 2012, 109 – 110).<sup>195</sup> Expounding further on Crusoe, McKenna continues:

But that is not to say that [Crusoe] is not a moral agent. It is not to say that we cannot truly say of him that he is cruel, or that his actions are morally bad or wrong, nor is it to say that he himself could not understand the judgment that his actions are cruel or bad. It is only to say that, as a moral-responsibility-idiot, one who is incapable of grasping and understanding the interpersonal exchanges central to a conversational theory, he is not blameworthy. (McKenna, 2012, 109 – 110)<sup>196</sup>

---

<sup>194</sup> Compare non-quality of will theorist, Marina Oshana’s (2004, esp. 269) take on these sorts of cases.

<sup>195</sup> The idea is that we presuppose that the individuals we hold responsible can respond to and issue certain types of reasons. See, e.g., Oshana (1997, 76 – 79; 2004, 269), Darwall (2006, esp. 8 – 10, 11 – 17, 58 – 60, & 65 – 90), and McKenna (2012, 80 – 88 & 105 – 110).

<sup>196</sup> Cf. Oshana (2004, 262 – 269).

For McKenna, then, one can be cruel and do wrong where their cruelty and wrongdoing fail to give us a reason to blame them. Thus, McKenna, along with many, holds that individuals like Crusoe stand outside the community of morally responsible agents.<sup>197</sup> The worry, if there is one, is that on the simple view I want to defend, folks like Crusoe count as being inside that community, since that view does not fully excuse or exempt them on the grounds that they lack the capacity to understand particular reactions some might have to their behavior. Furthermore, such an inability may often cause gaps in care.

McKenna-Type Conversational Theories seem to start at the wrong place. They begin by looking at the propriety of the act of *expressing* the attitudes that underlie instances of holding responsible and infer that their appropriateness is beholden to the same standards. That is a mistake. Even if McKenna is correct to hold that we gain our knowledge of whether an attitude is an instance of holding responsible or not by encountering expressions of it, the standard of appropriateness for having the underlying attitudes might remain distinct from the standards governing expressions of them.<sup>198</sup>

We should instead begin with what the attitudes being expressed are about and are

---

<sup>197</sup> Given parallels between Crusoe and Oshana's (2004) psychopath, many will see Crusoe as at least one type of philosophical psychopath. Some dissent from the common view that it is inappropriate to blame them, see, e.g., Talbert (2008; 2012). Cf. A. M. Smith (2008, 387 – 390). Shoemaker (forthcoming a, b) accommodates the dissent differently than I do here but our views are compatible.

<sup>198</sup> McKenna (2012, 69 n. 10 & 70 n. 11) appeals to de Sousa's (1987) idea regarding the way in which we come to understand our emotions to substantiate the idea that we should pay attention to expressions of blame to determine when blame-constituting attitudes are appropriate. The appeal is confused. In a nutshell, the appeal simply trades on a misunderstanding of de Sousa's theory of emotions. On de Sousa's view, roughly, the content of our emotions is fixed by engaging our social world. However, the appropriateness of one's emotion is not *wholly* answerable to whether the social situation calls for such an emotion. Rather, it is also answerable to whether the target of the emotion indeed is, say, something to be feared or in our case worthy of blame. [omitted]

sensitive to (i.e. what it is about the world that these attitudes track). For modern-day quality of will theory founder, P.F. Strawson, these attitudes track and represent the interpersonal significance that certain decisions have. Quality of will views pick up on this idea and hold that the quality reflected in an individual's action is determined by its moral significance. Thus, it is *not* one's ability to understand your praise/blame of her that renders her worthy of praise or blame. Rather, it is the cruelty or kindness manifested by what she *qua* agent does that makes her so worthy.

Elaborating a bit, Strawson tells us that he focuses on the reactive attitudes because what he wants “to insist on is the very great importance that we attach to the attitudes and intentions towards us of other human beings, and the great extent to which our personal feelings and reactions depend upon, or involve, our beliefs about these attitudes and intentions” (Strawson, 1962, 75).<sup>199</sup> Strawson's ultimate conclusion is that our moral responsibility practices remain justified and coherent in the face of determinism without appeal to the distribution of consequences we are likely to achieve by holding each other responsible. And, he reaches that conclusion, roughly, on the basis of his claim that the attitudes at the center of these practices do not respond to reasons grounded in theses like causal determinism or those grounded in facts merely about how best to manage individuals for personal or collective interest. Rather, for Strawson holding responsible is reasonable only in light of the quality of the will manifested and reflected in our actions. As Pamela Hieronymi points out, Strawson is at least in part able to reach this conclusion due to the fact that the reactive attitudes featured in blame

---

<sup>199</sup> Just prior to this, Strawson characterizes the reactive attitudes as “the non-detached attitudes and reactions of people directly involved in transactions with each other; [...] of such things as gratitude, resentment, forgiveness, love, and hurt feelings” (Strawson, 1962, 75).

“involve [...] a complex set of judgments or evaluations—a set which necessarily includes, as a unifying feature, the judgment that disregard was shown” (Hieronymi, 2004, 133). To put this in terms of our best scientific understanding of the types of attitudes to which Strawson directs our attention, the attitudes constituting instances of holding responsible represent *core relational themes* that track the interpersonal significance of wrongdoing.<sup>200</sup> Thus, the only considerations that render these attitudes inaccurate and, so, inappropriate are those that speak to that significance.<sup>201</sup>

The fact that Crusoe acts cruelly despite understanding that doing so is wrong is a fact that bears interpersonal significance. For instance, it helps us determine whether and why anyone should or shouldn't have certain sorts of relationships or dealings with Crusoe. Hence, our responding to Crusoe as morally responsible for his treatment of Rover seems completely appropriate. However, the fact that, were we to express that attitude, Crusoe would fail to understand our expression seems beside the point. Even if his inability to grasp such expressions itself carries interpersonal significance and so may legitimate other reactions that track such things or limit the ways we may outwardly treat Crusoe, it does not fully excuse or exempt him from attitudes that constitute blame.

Generally, *Crusoe-type cases*, as I will call them, essentially involve an individual who has intentionally flouted a moral norm but lacks a particular feature other than the capacity to respond to reasons. Since responsibility responses inform us about such

---

<sup>200</sup> Robert Lazarus (1991, 22, 65 – 83, 116 – 124, 165 – 166 & 193 – 213) seems to have coined the term ‘core relational themes’, which Jesse Prinz nicely glosses as characterizing “a relation that pertains to well-being” (Prinz, 2004, 15). See also, C. A. Smith & Lazarus (1991). Cf. Prinz (2004, esp. Chpt. 3), Greenspan (2004, esp. 207, 218 – 219 \*though she does not use this terminology), Griffiths (2004, esp. 243 – 247), Frijda (2011, esp. 79), Salovey *et al* (2011, esp. 537).

<sup>201</sup> Hieronymi, (2004, 135). Cf. Fisher & Manstead (2008) for empirical support regarding the Hieronymi's idea that these sorts of attitudes serve a social function.

individuals insofar as they are (in part) constituted by a judgment or recognition of will, Hieronymi seems to correctly assess the status of these individuals as members of the community of morally responsible agents. The judgment that Crusoe-type individuals have shown ill will or an insufficiently good will is warranted, since it is only facts about the quality of her manifested will that determines that warrant. In other words, it is only those facts that “bear on whether the content of the judgment is correct” (Hieronymi, 2004, 126). Thus, I see no reason to think that individuals like Crusoe cannot manifest wills of no interpersonal significance. Thus, their inability to understand the communicative import of responsibility responses seems not to bear on whether Crusoe-Type individuals are fit for negative moral appraisal of their cruelty. And, so, if an absence of care has its causal history grounded by their inability to understand blame-constituting attitudes, where correcting that absence is within the reach of their agential capacities, the injuries they cause through negligence is, all things equal, worthy of appraisal.

### 3: CONCLUSION

Negligent injury involves careless oversight or unreasonable lack of circumspection. Contrary to concerns about how we should understand the manifests relation, quality of will views not only remain viable, they also make good sense of negligent injury. The basic idea is that the presence of an absence under one’s agential control can color the quality of their will. Furthermore, the view that some unreasonable absences of care or regard are due to an inability to understand praise-/blame-constituting attitudes does not

itself speak to whether individuals capable of manifesting a will via their action are fully excused or exempt from moral appraisal. Indeed, supporting their absence from the community of morally responsible agents on the basis of that communicative lack is misguided. Certainly some gaps and momentary lapses in reason may be beyond the reach of an individual's agency. And, some may also be caused by a failure to understand expressions of responsibility responses. Still, where more sensitive reasoning could have prevented harm, cruelty, or disrespectful behavior the manifestation of will reflected is still morally appraisable. Thus, the presence of absence in manifested will is not itself a reason to withdraw praise or blame. Rather, an otherwise well-intentioned, magnanimous individual's harm causing negligence is, all things equal, fit for our resentment.

One interesting way to resist my conclusion from the perspective of quality of will theorists is to object on the grounds that I draw the boundaries of the moral community too widely. The objection might find traction in Strawson's program. More importantly, the objection might occlude a Strawsonian argument against the skeptic like the one I hope to offer. To press this sort of concern one argues that my account entails that we have rational grounds for responding in ways constitutive of holding responsible to the behavior of those we take to be natural, mechanistic things. But, the thought goes on, the claim that we do not have such grounds is a crucial piece to Strawson's argument.

In the next chapter I distinguish between viewing something as a natural, mechanistic thing and viewing it as a *mere* natural, mechanistic thing. On my account, those merely natural, mechanistic things are rightly regarded as outside the moral

community. I show that this salvages the path I aim to take in response to the skeptic. It also further develops my view about the wideness of the moral community.

## VII: OPTIMISM AND RESENTMENT

We react to the behavior of those around us with affection, gratitude, love, resentment, indignation, sadness and other sentiments. Whether we always express them or not, we typically feel complex emotions when, say, a close friend forgets our birthday or joins in mocking us as well as when someone exploits the poor position of another, maliciously assaults their kin or murders in ‘cold blood’. Likewise, apparently appropriate feelings are often provoked by acts of kindness, gentle smiles from intimates, thoughtful support from a colleague, and demonstrations of compassion. P. F. Strawson surmised that it would be, as a practical matter, inconceivable to completely jettison these sorts of reactive attitudes.<sup>202</sup> For Strawson, this would involve deciding to exclusively adopt the broad psychological orientation which colors individuals as natural, mechanistic things—i.e., the so called *objective stance*.<sup>203</sup> Such exclusive adoption amounts to a saturation of experience that affectively causes the adoptee to always encounter everyone everywhere as natural and mechanistic however else she might understand them to be and, hence,

---

<sup>202</sup> Strawson (1962). Cf. Beardsley (1960, esp. 5 – 6). For excellent collections on Strawson’s essay see, Russell & McKenna (2008) and Shoemaker & Tognazzini (2015). Strawson’s set of reactive attitudes is large and includes at least those cited at the beginning of the paper, i.e. attitudes like resentment, indignation, guilt, gratitude, love, and hurt feelings. Some like Wallace (1994, 25 – 40; 2011) and McKenna (2012, 65 n. 7) accept a narrower set. This will not affect my argument.

<sup>203</sup> Though Strawson opts to speak in terms of adopting an *attitude* instead using stance language, I agree with Coleen Macnamara that it “accurately captures Strawson’s view and adds clarity” (2011, 83 n. 1). Importantly, a stance *should not* be confused with a belief. As Coleen Macnamara explains, stances structure our “presumptive interpretations [...] behavioral dispositions [...] and [...] emotional proclivities” (2011, 84). The exclusivity here is about whether one takes the stance toward all things at all times or not. It is not about whether the stance is *the only* stance she takes toward anything. Also, Strawsonians are not clear what they take *adopting* a stance to be. This does not affect the arguments here. [omitted]. The thoughtful commentary of two anonymous referees encouraged me to move this note forward in the text, rearrange it, and add a bit about the exclusivity in question and be more clear about Strawson’s choice of terminology.

ostensibly, to encounter them as unfit for the reactive attitudes which Strawson thinks serve as a key basis for deep, interpersonal relationships. Though Strawson-style accounts of moral responsibility differ, they retain Strawson's idea that exclusively adopting the objective stance is discordant with deeply personal relationships.<sup>204</sup> Call that idea, *Objective-Stance Pessimism (OS-Pessimism)*. This paper rejects that keystone pessimism, shows that Strawson-style accounts can do without it, and thereby sketches a superior non-skeptical Strawsonian picture of moral responsibility.

Interestingly, OS-Pessimism is a point of agreement between many who accept and many who reject the appropriateness of the reactive attitudes (i.e. responding to individuals as morally responsible). Generally, OS-Pessimism is meant to show affinity for the suspicion that being psychologically oriented toward everyone as natural and mechanistic profoundly affects your interpersonal relationships with them.<sup>205</sup> Even skeptics about moral responsibility who reject OS-Pessimism because they are more sanguine about retaining deeply personal relationships in the face of exclusive adoption of the objective stance agree that the saturation of experience that such adoption induces profoundly changes those relationships given that it undermines the propriety of the reactive attitudes (Derk Pereboom (2001; 2014), Tamlar Sommers (2007)).<sup>206</sup>

In a nutshell, against this tradition and contrary to those who join me in abandoning OS-Pessimism, I take my rejection of it to do two things. First, it basically

---

<sup>204</sup> Strawson (1962, esp. 79 & 81 – 82). Cf. Bennett (1980, esp. 40), Wolf (1981), Benson (1990), Wallace (1994, 28 – 31), Pettit (2001, esp. 66 – 70), McKenna (2012, esp. 64 – 78 & 211 – 213), and Shabo (2012).

<sup>205</sup> Strawson (1962, 82). Cf. Sommers (2007), Shabo (2012), Pereboom (2001, Chpt. 7; 2014, Chpts. 6 & 8).

<sup>206</sup> See, e.g., Pereboom (2001, Chpt. 7; 2014, Chpts. 6 & 8) and Sommers (2007). Roughly, the idea is that personal relationships do rest solely on the fitness of reactive attitudes. So, even though causal determinism undermines moral responsibility by rendering the reactive attitudes inappropriate, it does not undermine all of the attitudes that serve to support deeply personal relationships.

rids Strawson-style accounts of moral responsibility of a central but false thesis. Second, rejecting OS-Pessimism cancels a significant concession that many non-skeptics make to skeptics—namely, that living under the assumed truth that individuals are natural and mechanistic would radically change our practical lives. I go on to reinforce non-skeptical Strawson-style accounts by putting to work their underemployed idea that the extent to which an individual is poised to respond to and on behalf of others as morally responsible creatures depends primarily on the extent to which she is oriented toward them as participants in the moral community. It follows from my rejection of OS-Pessimism that, *pace* tradition, whether an individual encounters someone as a participant in that community does *not* depend on whether she is oriented toward them as natural and mechanistic. Thus, I show the need for an improved groundwork for non-skeptical Strawson-style accounts of moral responsibility and I fill that need by drawing upon traditional Strawsonian insights.<sup>207</sup>

More precisely, then, the traditional pessimism about retaining a basis for morally important attitudes in the face of exclusively adopting the objective stance implicitly relies on the

*No-Difference Thesis:*            There is no difference between seeing something as a natural, mechanistic thing and seeing something as a *mere* natural, mechanistic thing.<sup>208</sup>

I argue that the no-difference thesis is false. I thereby reject OS-Pessimism and secure *objective-stance optimism* (OS-Optimism) (§1). I follow by refashioning the crucial link

---

<sup>207</sup> I thank three anonymous referees that reviewed this paper for pushing me to make the stakes clearer.

<sup>208</sup> *Seeing something as* is taken here to be distinct from *believing that something is*. Rather, it is *being psychologically oriented toward something as*. Also, here foreword I sometimes drop ‘exclusively’ for ease of prose.

Strawson forged between the objective stance and the attitudes involved in deeply personal and moral relationships by leaning on traditional Strawsonian views about interpersonal relations amongst those who encounter each other as participants in the community of morally responsible agents (§2). The upshot is that OS-Optimism can and should replace OS-Pessimism as part of the foundation for our Strawson-style accounts of being and holding responsible. If I am correct, the common ground between non-skeptical Strawson-style accounts of moral responsibility and global skepticism about it is shattered as is a key thesis that helps the latter support their skepticism. I begin to lay this more firm foundation by showing that we have no reason to accept the no-difference thesis.

#### 1: SECURING OS-OPTIMISM

OS-Optimism is primarily secured by the fact that OS-Pessimism is fundamentally mistaken. I work to show this in three ways. First, I argue that Susan Wolf (1981) and Seth Shabo (2012) offer arguments for OS-Pessimism which depend on the no-difference thesis. I argue that the no-difference thesis is false and follow by showing that the mistake of leaning on it pervades OS-Pessimism.

Second, then, I return to Wolf but focus on her broader view. According to Wolf's broader view, even if our moral concepts and practices are not completely diminished in the face of determinism, there is a deep paucity of meaning behind our attitudes, actions, and projects in a deterministic world. Some might rely on this sort of thought to motivate OS-Pessimism in a manner that might circumvent reliance on the no-

difference thesis. However, developing Nomy Arpaly's (2006) rejection of Wolf's view, I show that the latter is nevertheless backed by the no-difference thesis.

Finally, I close this section by turning to Strawson's general defense of holding responsible, which is taken to be the source of OS-Pessimism. Close analysis of Strawson's overall defense reveals that it does *not* require the no-difference thesis, and, so, does not require OS-Pessimism. This does more than simply locate possible error at the source; it steps toward undercutting Strawsonian motivation for OS-Pessimism. In the section that follows, I draw upon the Strawsonian idea that the extent to which we respond to individuals as responsible depends on the extent to which we are oriented toward them as participants in the moral community to do the work that OS-Pessimism did. The upshot is that OS-*Optimism* is secure and non-skeptical Strawson-style accounts are strengthened by an adoption of it.

### 1.1: CARING PERSONALLY ABOUT NATURAL THINGS

Wolf and Shabo offer structurally similar arguments in favor of OS-Pessimism. After attempting to identify what it is to view something from the objective stance, each claims that we cannot be involved in deeply personal relationships with things toward which we are so oriented. Basically, I argue that we can agree with each regarding what is necessary for being in a personal relationship and consistently maintain that these elements need not be absent from a relationship where one adopts the objective stance. So, in defense of OS-Optimism, for the sake of argument, I grant their views about the conditions for being involved in deeply personal relationships, while leaning on the point

that there *is* a crucial difference between seeing something as a natural, mechanistic thing and seeing something as a *mere* natural, mechanistic thing.<sup>209</sup> Consider Wolf first.

Wolf begins by pointing out that, in a world in which everyone has universally adopted the objective stance, people would still imprison, praise, applaud, criticize, thank, and shame as they saw fit to do.<sup>210</sup> According to Wolf, these reactions would simply express a desire that a behavior or attitude persist or discontinue.<sup>211</sup> Wolf sees this objective-stance world as significantly lacking. She holds that in such a world one can choose to interact with others *only* on the basis of “whether they offer, to a sufficient degree, the proper combination of pleasure and practicality” (Wolf, 1981, 391). On Wolf’s view, we are thus restricted to choosing our friends and lovers “as we now choose clothing or home furnishings or hobbies” (Ibid.). While this might allow individuals to form “strong attachments to their cars, their pianos, [...and] their pets” (Ibid.), Wolf holds that one cannot form deeply personal relationships on such a limited basis. I disagree.

We should distinguish between seeing something as a natural thing and seeing it as a *mere* natural thing.<sup>212</sup> For instance, I might see the flower pot, the ant, and my brother, Hugh, each as natural, mechanistic things insofar as certain outputs follow certain inputs—that’s ‘how they work’, as it were. However, *contra* Wolf, I might choose to interact with Hugh and other things like him in a manner that takes into account

---

<sup>209</sup> This use of ‘mere’ does quite a bit of work throughout. I hesitate, however, to offer a general description of the difference it marks in this context, since the authors upon which I focus have slightly different ideas about what it is to see something as a natural, mechanistic thing. Still, my discussion should make clear the distinction. I thank [omitted] for suggesting this footnote.

<sup>210</sup> Wolf (1981, 390 – 391). Cf. Strawson (1962, 79) & Watson (1987, 225).

<sup>211</sup> Wolf (1981, 390 – 391).

<sup>212</sup> See footnote 6.

*more than* how pleasing and practical such interaction would be. Tamler Sommers makes this point as well<sup>213</sup>:

When you take the objective attitude towards other human beings, you do nothing more than see them as natural things. But a human being is still a human being—the most exciting, infuriating, unpredictable, lovable, loathsome natural thing in the world. So when we adopt the objective attitude, we would not merely find people useful or amusing. We would not choose our friends as we would choose home furnishings, hobbies, songs, pianos or pets. [...] We choose friends as we choose human friends—that is all. Nothing in the objective attitude prevents us from recognizing, appreciating, *cherishing* the rich and wonderful qualities of another person. It remains the choice that brings the greatest rewards and the deepest disappointments in all of human existence. (2007, 326, his emphasis)<sup>214</sup>

We see, then, that, in her description of the objective stance, Wolf seemingly wrongly assumes that engaging with individuals as natural things precludes choosing to interact on a basis that goes beyond how pleasing and practical such interaction would be.

On Wolf's understanding, if an individual sees the world as natural and mechanistic, then she is and feels released from the strains of commitment that allow her or bind her to regard certain individuals in a manner that recognizes their special nature. But, that seems not to be the case. An individual's recognized status as natural and mechanistic leaves completely undetermined whether we should feel gratitude toward her when she offers a helping hand or feel indignation toward her where she exploits another for profit. I may find reasons that help me loosen the grip of commitment that in part grounds the fitness of the love, adoration, and respect that I feel toward my brother in

---

<sup>213</sup> Shabo (2012, 138) also recognizes this as Sommers's point. It is not clear, however, that Sommers sees the insight to which I draw attention, since he takes adopting the objective stance to commit one to seeing others as not deserving of the reactive attitudes. See, Sommers (2007, 324) and §2 below.

<sup>214</sup> Cf. Pereboom (2001, 199 – 207; 2009; 2014, Chpt. 6) allows these relationships but holds that they do not involve fitting reactive attitudes. See [omitted]

order to guide him through a difficult decision or relay a harsh truth. Furthermore, the reasons that aid in loosening that grip need not have anything to do with Hugh's being natural and mechanistic. Indeed, it seems I may already be oriented toward him as such and still be searching for a way to escape the authority of my commitment to him and all persons. Alternatively, should I find myself unmoved by his offer to take on the burden of caring for our aging parents or his locating a serious job opportunity after I lose my own due to reorganization in the company, this seems due to attitudes I have toward him beyond any stance I take regarding his metaphysical nature as less than supernatural or unable to transcend causal laws. It seems that I balk at responding to him in a manner that rests on something more than the degree to which he offers a particular combination of pleasure and practicality only where I have adopted other attitudes about him, our world, or being natural and these attitudes do the work of impeding my responsibility responses. Hence, I may see Hugh as natural and mechanistic, and yet not merely so.

In short, then, Wolf's mistake is that she assumes the truth of the no-difference thesis. The mistaken assumption plagues OS-Pessimism. To get a sense of this, consider Shabo's argument.

Concentrating on the reasons why we experience particular emotional reactions to offensive behaviors and attitudes, Shabo adds to Wolf's account. Shabo notes that, in response to a slight, one "may become emotionally withdrawn, feel angry, hurt, or shocked, seek opportunities to retaliate, demand an explanation, and so on" (Shabo, 2012, 139). According to Shabo, in some cases this is due to the fact that it matters "a great deal to her *why* the slight occurred, where this is precisely a question about what the offending behaviour reveals about the other's attitude towards her" (Ibid.). And, he

continues, the reactor in this case “*cares in an essentially personal way* about the other’s treatment of her” (Ibid., his emphasis). On Shabo’s view, we react like this when we care about another’s behavior “*independently* of what it presages about his future conduct, or how that conduct might affect one’s (or a third party’s) interests” (Ibid., his emphasis). According to Shabo, exclusive adoption of the objective stance is incompatible with this sort of care.<sup>215</sup>

Contrary to Shabo, I see no reason to think that I cannot coherently question what the behavior of another reveals about her attitudes toward me while maintaining the broad psychological orientation that paints her as a natural thing, i.e., while exclusively maintaining the objective stance. Suppose, for instance, that I interact with and begin to take notice of a particular human named, *Gena*. My interest is piqued by her bright smile, cheerful, warm, and generous character under toned with deep intelligence, religious curiosity, and an attentiveness to life typically due to experiencing and emerging anew from great personal tragedy. Suddenly, I’m hooked. Imagine further that she is likewise drawn to me. Eventually, we move from exchanging curious glances to spending time together at social gatherings and alone on dates. We laugh with each other and learn about the personal struggles and triumphs, hopes and interests, reservations and ambitions we each possess. As we move forward, I see her as special and know that she sees me as special too. Thus, I expect that Gena will do certain things and hold certain attitudes. For example, I expect most normal individuals to typically take what they view as the best means to their ends when they can, and I expect no less from Gena. But, I

---

<sup>215</sup> Shabo understands *taking something personally* as the manifestation of a disposition *to take something personally* which is “partly constitutive of caring in this essentially personal way about that person’s behaviour toward us” (2012, 139).

don't expect just anyone to see something at a store, buy it, and offer it to me thinking I'd enjoy it. Likewise, Gena doesn't expect others to cheer her up with a short note sneaked into her lunch bag or a dozen Gerber daisies. However, she and I do expect those sorts of behaviors from each other. Of course, Gena also expects that I won't do things intending to upset her, that I won't mock her behind her back or seek romantic comfort in the arms of another, and I expect the same of her. We generally value the attitudes of, say, mutual regard and respect we take to be expressed by these behaviors.<sup>216</sup> And, in particular, we value them as the attitudes expressed toward us by each other. Suppose further that I am oriented toward all things including Gena in such a way that all are seen as natural and mechanistic. Is this story suddenly rendered incoherent or implausible? I don't think so.

It is perhaps clear that, were I to view Gena as a *mere* mechanistic thing, it would be odd if I also saw in her behavior expressions of regard for me and others. However, I see no reason to think that I must refrain from viewing her as a natural, mechanistic thing in order to see many of her behaviors in this light. Seemingly, for example, whether my behavior generally follows the pattern Gena expects and demands is important to her in the sense that she cares about it *independently* of what it portends about her future interests.<sup>217</sup> Thus, it seems that the relationship just described has what Shabo deems necessary to count as a deeply personal one, for it is a relationship holding between two

---

<sup>216</sup> I am indebted to Wallace's (2011) development of his Strawsonian view.

<sup>217</sup> I imagine that enough deviance on my behalf from a pattern of behavior Gena takes to evince the mutual regard and respect she values would likely lead her to avoid me. And, it's quite plausible that cessation in light of sufficient deviation is connected to what my deviation reveals about my attitudes towards her. Ostensibly, this is accompanied by various emotions. Below I argue that feeling this does not betray the objective stance. I omit those emotions in my example only to forestall this discussion until the next section. Note also that I sketched this as Wallace (2011, 366 – 372) does. Filling this out further along his lines: Gena generally values the attitudes of, say, mutual regard and respect, and via our interaction has come to value my attitudes of this sort in particular. Although the Wallace sketch is in terms of expectations and demands, nothing is lost here if we characterize this all in terms of desire(s).

individuals whose behavior affects them beyond a mere concern for how they might manipulate each other to gain personally or benefit society generally. Perhaps, these individuals view each other as natural, mechanistic things. Perhaps, they don't. But, importantly, it seems that nothing about the coherence of this story hinges on whether they do. Indeed, it seems that I could not excuse bad treatment of Gena or be unhitched from a commitment to her purely on the basis of whether I engage her as something that responds in specific ways when presented with certain actions, attitudes, and circumstances. And, were Gena to respond to my offering such an excuse for manipulating her by asking whether I see her as a mere mechanism and means to my own or society's end, it seems we should understand her to have made a sensible and undermining quandary about the legitimacy of the appeal to her nature as a certain sort of mechanism explicit in my excuse. "Yes I am," Gena might say, "the sort that cares about you." I see no reason to see her rebuttal as anything but genuine and appropriate.

Thus, whether one embraces a stance that colors individuals as natural and mechanistic she could remain interpersonally bound to them and herself in a manner that renders it fitting for her to feel love, admiration, gratitude, pride, resentment, indignation, righteous anger, or guilt toward them or herself. Wolf's argument and Shabo's argument rest on the false assumption that such is not the case. And, the culprit here appears to be the mistaken view that there is no difference between being oriented toward something as natural and mechanistic and being oriented toward something as *merely* natural and mechanistic—i.e., the no-difference thesis.

Still, there is more to Wolf's influential paper than the argument just canvassed. In an effort to push back against my critique, one might pick up on Wolf's general theme

that even if our moral concepts and practices are not completely diminished in the face of determinism, there is a deep paucity of meaning behind our attitudes, actions, and projects in a deterministic world. This involves arguing for OS-Pessimism by employing a parallel view according to which an individual encounters the emptiness Wolf cites due to the fact that their adoption of the objective stance has inundated their experience. The basic idea is that this individual now engages everything in the world in terms of how it works as a natural operating system whereas prior to adopting the objective stance she experienced some of these things as acting for reasons. This route is different than the one's delineated above, since it intentionally seeks to motivate OS-Pessimism in a manner that might make sense of the no-difference thesis without relying on it. I turn, then, to consider this alternative approach, show that it nevertheless rests on the no-difference thesis, and offer further reason to reject that thesis in a manner that further develops my optimistic vision.

## 1.2: THE ROBOT CLAIM

Attempting to elicit an intuition in favor of her version of OS-Pessimism, Wolf describes the case of a sophisticated robot. She writes:

Let us imagine an individual who has been and continues to be very completely and elaborately programmed. He is programmed not only to make various choices and perform various actions, but also to engage in various thought processes, to form various second-order volitions and so forth, in coming to perform these actions. Indeed, this individual is programmed in such a way as to appear to be an ordinary human being in every respect. If no one were informed that this individual was programmed, he would appear both to us and to himself to be 'one of us'. [...] One might imagine the relation between robot and programmer to be very much like a possible relation between author and

character; [...] though the robot might choose to perform the actions he performs, he chooses to perform them only because he is programmed to so choose. Though his decisions and judgments may be preceded by thoughts which look or sound like reasons, he cannot be said to reason to these conclusions in the way we do. [...] (Wolf, 1981, 395 – 396)<sup>218</sup>

Wolf holds that, if we knew that this was the true story of our ‘friend’ Robot, “the only way of living in accordance with the facts would be by regarding the robot solely with the objective attitude” (Ibid., 396, footnote omitted). We would, Wolf thinks, have to hold that “the robot is not a free and responsible being” (Ibid.). Putting this together with her thought that deeply personal relationships are incompatible with the objective stance, Wolf’s thought could be employed to support OS-Pessimism in a manner that appears to circumvent reliance on the no-difference thesis. In this vein, contrary to my contention, one takes the line that the fitness *and so the value or meaning of* the attitudes the recognition of which provides the basis for deeply personal relationships is completely depleted where they are the attitudes of a natural, mechanistic being like the robot Wolf describes. A proponent of this line avers that the depletion is due to the fact that the robot’s behavior is caused not by the *content* of the program but merely by whatever *delivers* that content.<sup>219</sup> Thus, the thought continues, always living under the assumption that everything everywhere is like the robot natural and mechanistic where you once

---

<sup>218</sup> The main parts that I leave out of Wolf’s original description suggest that the robot is being *manipulated* by someone else. But, this confuses the main issue, since manipulation cases inherit worries that are arguably separate from worries concerning whether OS-Optimism is true. See, Pereboom (2001, 112 – 117; 2014, Chpt. 5) for his influential ‘Four-Case Manipulation Argument’ and catalog of responses to it. [omitted]

<sup>219</sup> More charitably, perhaps, Wolf is saying that the freedom and responsibility conferring features of her mental content are undermined. This may allow Wolf to avoid the less discriminate version of Arpaly’s argument described below, but not my ultimate response. I thank [omitted] for discussion on this point and for the friendly suggestion.

might have taken and individual's behavior to be backed by attitudes of mature friendship or reciprocal love, you encounter them as backed by nothing of that sort. Rather, you experience them as void of content in this regard. It is this line of thought that might offer the OS-Pessimist an alternative path to her pessimism that avoids my concern that the view rests on maintaining the no-difference thesis.

To help fill out this alternative path, following Arpaly (2006), we can characterize the dispute between OS-Optimism and OS-Pessimism as a dispute over<sup>220</sup>:

*The Robot Claim:* Being responsive to the content of moral, artistic, romantic, and otherwise important reasons is discordant with being a natural, mechanistic thing.

On this construal of the debate, *OS-Pessimists* can accommodate the view that the robot claim is true (although distinct from their pessimism), while *OS-Optimists* cannot.<sup>221</sup> It is easy to see how the robot claim could motivate OS-Pessimism. Following Arpaly's depiction of Wolf further, the idea is that where we are natural and mechanistic

Artistic beauty, the admirable attributes of one's beloved, the despicable nature of selling divine absolution, or the moral wrongfulness of lying might cause us to do things, but only in the way that an onion can cause tears or Maria Callas's voice can cause a glass to break: in a meaningless, robot-like way. (Arpaly, 2006, 45)

The idea is that human beings' mental states and actions can be caused "by things that cannot determine the states and deeds of robots: desires, beliefs, artistic inspiration, love, religious faith, a conscience that disapproves of lying, awareness of artistic beauty, the admirable attributes of one's beloved, the despicable nature of the practice of selling

---

<sup>220</sup> Arpaly (2006, 45).

<sup>221</sup> The fact that Strawson was not interested in reasons-responsiveness does *not* preclude Strawsonians attracted to OS-Pessimism from finding a path via the robot claim that motivates their view. [omitted]

divine absolution, or the moral wrongfulness of lying” (Ibid.). Ostensibly, this idea supports the view that were we broadly, psychologically oriented toward ourselves and others in such a way that we and they were painted as natural, mechanistic things we would be mistaken if we also saw the behavior of ourselves and others as expressive of, e.g., regard, disrespect, or esteem.<sup>222</sup> Our mistake, the OS-Pessimist maintains, would be akin to thinking that it was, at least in part, say, the sorrowful meaning of the words in Maria Callas’s song that caused the glass to shatter.<sup>223</sup>

For instance, the OS-Pessimist might claim, roughly, that the individuals toward whom we take the objective stance must be seen by us as individuals who cannot and do not respond to the content of reasons.<sup>224</sup> Thus, the thought goes on, those toward whom we take the objective stance are viewed in a manner that colors their behavior and attitudes as caused by something other than the content of, say, an expression of love or friendship. On that basis, the OS-Pessimist concludes that their view is true, since (I) being in a deeply personal relationship requires engaging others as beings who respond to the content of expressions of romantic love or mature friendship, which (II) is simply *not* something that one who has exclusively adopted the objective stance does because (III) she is oriented toward them in a manner that colors them as beings that are simply pushed about, beings, that is, who do not do things in response to the relevant content.

At least two things should be said in response to this line of reasoning. First, whether OS-Pessimism’s accommodation of the robot claim actually motivates the view or not, Arpaly offers good reason to think that the robot claim is highly implausible. As

---

<sup>222</sup> Expressive, that is, of what Strawson took the reactive attitudes to track—namely, qualities of will. See below. [omitted]

<sup>223</sup> Arpaly (2006, 45 – 46). See footnote 221.

<sup>224</sup> Or, the personal-relationship conferring features of this content. See footnote 221.

she points out, defenders of the robot claim are forced to hold that, in a deterministic world, content is *not* causally efficacious.<sup>225</sup> But, seemingly, the truth or falsity of determinism should have *no bearing* on whether content is causally efficacious.

Explaining content efficacy, Arpaly writes:

Imagine an opera singer singing about love. She hits a high E; as a result, a glass breaks. In this case a love song caused the glass to break but, presumably, the breaking of the glass had nothing to do with the content of the song. If the song had been about, say, frogs, the glass would have broken just the same, so long as a high E was hit at the requisite volume. In other words, though the song happened to have a certain content, the content of the song was not efficacious in breaking the glass. If, on the other hand, the song were to send a recently divorced listener into tears, the content of the song would seem to have played an important causal role in this latter incident. (Arpaly, 2006, 45 – 46, footnote omitted)

As far as I can tell, Arpaly is correct that whether causal determinism is true should have no bearing on whether various types of causation can occur. Notice, the former is just a view on which, roughly, whatever *can* cause events in this world the same pre-circumstance and the same laws of nature will generate the same post-circumstance. Thus, it would be astonishing if content efficacy required *indeterminism*. And, since the incompatibility of content efficacy and determinism is central to touting the robot claim, an argument built upon it should be rejected.<sup>226</sup>

---

<sup>225</sup> See footnote 221.

<sup>226</sup> There are, of course, responses available to the OS-Pessimist. For instance, they might try adjusting the robot claim to allow for content efficacy but rule out, say, reasons-responsiveness. Arpaly (2006, 49 – 65) suggests this response and variations of it. Rather than reproduce the dialogue, I simply cite my general agreement insofar as her main point is concerned. Arpaly's main point is that standard ways of proceeding here rely on the view that causal determinism rules out *something* vitally important for personal relationships and/or 'on-one's-own'-creativity, but neither side can substantiate this claim. As Arpaly makes clear, each attempt to substantiate runs into serious trouble. Either they make the astonishing claim that whether content is causally efficacious depends on the truth of causal determinism or they allow

Second, put aside for a moment the fact that the robot claim seems to entail a false thesis concerning the compatibility of determinism and the causal efficacy of content. The argument for OS-Pessimism constructed by appealing to the fact that the view accommodates the robot claim is essentially the same as those previously critiqued insofar as it actually does rely on the no-difference thesis. The main difference here is that (a) one cites as a requirement for being involved in deeply personal relationships that the parties must understand each other as beings responsive to the content of reasons and (b) follows by claiming that adoption of the objective stance precludes seeing each other as such. But, (b) rests on the problematic assumption that viewing something as a natural, mechanistic thing is to view it as a *mere* natural, mechanistic thing. Here, the idea is that the latter but not the former precludes choosing to interact with an individual on the basis that they respond to, say, the content of reasons for romantic love or mature friendship. Thus, though the appeal to the robot claim offers an alternative, non-Strawsonian route to OS-Pessimism, it is a route that nevertheless still relies upon the no-difference thesis. Furthermore, in light of Arpaly, we have reason to doubt the veracity of this position, since it suggests that once one adopts the objective stance toward something she has adopted an orientation toward it that paints it as a thing that cannot respond to the content of reasons. However, if Arpaly is correct, then adopting the objective stance toward others should have no bearing on whether I construe them as beings able to respond to the content of reasons. For one, to do so assumes that being natural and mechanistic is to be causally determined. And, for another, even if that were

---

content efficacy in a deterministic world. But, the latter is enough for reasons-responsiveness, and, so, it is enough for deeply personal relationships and ‘on-one’s-own’-creativity. Arpaly (2006, 65 – 85) follows by putting forward her own view of reasons-responsiveness. See also Arpaly (2003, esp. 125 – 131). I remain neutral here about the correctness of her account of reasons-responsiveness.

the case, the truth of determinism should have no sway in the debate over content efficacy. Therefore, this robot-claim path to OS-Pessimism is no better than those canvassed so far.

Perhaps, however, one can take a slightly different avenue. Consider, for instance, a path to OS-Pessimism that picks up on Wolf's thought that the significance of many of our attitudes and practices is lost in the absence of free will. On this path, we make use of the popular thought that one must see oneself as free in order to see oneself as answerable to practical demands or able to effectively deliberate about what to do. On this more direct route to OS-Pessimism, the idea is that adopting the objective stance toward oneself is incompatible with seeing oneself as answerable to practical demands or able to effectively deliberate about what to do, which is, in turn, required for entering into and maintaining deeply personal relationships.<sup>227</sup>

Again, at least two things should be said. First, the correct understanding of the term 'free' in the adage to which the alternate route appeals is up for debate. Must I see myself as having robust alternative possibilities in order to see myself as answerable to practical demands? Or do we simply have to see ourselves as the *source* of our actions and attitudes? Alternatively, do we simply have to see ourselves as able to be wholeheartedly committed to second-order desires that are aligned with our first-order desires? Or do we simply need to see ourselves as possessing some sort of guidance control? These are vexed questions that require deep theoretical and empirical work to help answer.

Importantly, we do not need to answer those questions in any particular manner to

---

<sup>227</sup> [omitted]

respond adequately to the argument under consideration, and this brings me to my second point. I see no reason for denying the claim that so long as we see ourselves as natural, mechanistic things and not *mere* natural, mechanistic things we can maintain that we are answerable to practical demands and able to effectively deliberate about what to do. That is, I see no reason for thinking that we cannot see ourselves as answerable to practical demands so long as we see ourselves as beings moved not *merely* in the manner that a glass is moved to break when in the presence of a high E *but also* in the manner that the newly divorced individual is moved to cry at the touch of the sorrowful content of a song being sung. So, I may see my brother, Hugh, and Gena as a natural mechanisms that output attitudes and behaviors given a certain input and but that alone doesn't preclude me from encountering them as beings that is stirred by the content of reasons. Perhaps spelling it out in these terms rings less romantic, but that is beside the point. The point is that OS-Pessimism simply implicitly relies on the no-difference thesis. Once we highlight reasons to think that there *is* a distinction between seeing something as a natural thing and seeing something as a *mere* natural thing we find that OS-Pessimism harbors illegitimacy at its core. Interestingly, it was perhaps there from the beginning.

### 1.3: THE SOURCE

Consider the explanation in Strawson that seems to spur towards OS-Pessimism<sup>228</sup>:

To adopt the objective attitude to another human being is to see him, perhaps, as an object of social policy; as a subject for what, in a wide range of sense, might be called treatment; as something certainly to be taken account, perhaps precautionary account, of; to be managed or handled or cured or trained; perhaps

---

<sup>228</sup> I must stress that I do *not* think it is clear that Strawson adopted OS-Pessimism.

simply to be avoided, though *this* gerundive is not peculiar to cases of objectivity of attitude. The objective attitude may be emotionally toned in many ways, but not in all ways: it may include repulsion or fear, it may include pity or even love, though not all kinds of love. But it cannot include the range of reactive feelings and attitudes which belong to involvement or participation with others in interpersonal human relationships; it cannot include resentment, gratitude, forgiveness, anger, or the sort of love which two adults can sometimes be said to feel reciprocally, for each other. *If your attitude towards someone is wholly objective, then though you may fight him, even negotiate with him, you cannot reason with him. You can at most pretend to quarrel, or to reason, with him.* (Strawson, 1962, 79, his early emphasis, my emphasis at the end, grammatical error corrected)

In light of the forgoing discussion, it should be easy to locate the no-difference thesis.<sup>229</sup> And, I have been arguing that we have no good reason to think that we cannot reason with those toward whom we adopt the objective stance, even if we have good reason to think that we are not able to reason with those we see as *mere* natural things. What has gone unnoticed, however, is that Strawson does not need to rely on the no-difference thesis to get his argument to go through.

To see this, note first Strawson's general argument. Strawson's main contention is that the attitudes involved in responding to individuals as morally responsible and those that form the basis for deeply personal relationships (i.e. the *reactive attitudes*), are, as a practical matter, inconceivable to do without. And, he goes on, to the extent that we *do* suspend them, such suspension answers only to considerations like whether ill will was manifested or whether we need "relief from the strains of involvement" (Ibid., 82).

---

<sup>229</sup> I should stress that contemporary discussions regarding what I call OS-Pessimism and OS-Optimism take the difference between seeing someone as capable of being reasoned with and seeing someone as unresponsive to such engagement to correspond to viewing them as natural, mechanistic or causally determined things. See, e.g., Stern (1974), Wolf (1981), Watson (1987), Sommers (2007), Shabo (2012), and McKenna (2012, 9 – 14 & 74 – 78). It is unclear whether this correspondence runs true to Strawson.

Further, for Strawson, these considerations have nothing to do with the truth of causal determinism. On (roughly) these premises, Strawson concludes that the coherence and justification of our moral practices does not depend on the truth or falsity of causal determinism. In support of his claim that the truth of causal determinism is not a consideration to which we attend in order to determine whether a sufficiently good will is being shown, Strawson brings in talk of the objective stance.

Strawson attempts to support his claim by elucidating two general cases in which we take it that regard is not being shown. He divides these cases broadly as follows<sup>230</sup>:

*Divorced-Will Cases:* C is a divorced-will case if and only if responding with any of the reactive attitudes is inappropriate due to the fact that the behavior toward which they are directed is not a manifestation of will because the perpetrator's will *is not* part of the cause of the behavior.

*Neutral-Will Cases:* C is a neutral-will case if and only if responding with any of the reactive attitudes is inappropriate due to the fact that the behavior toward which they are directed *is* the manifestation of will, *but* the agent was incapacitated in some way.

Strawson goes on to identify two subgroups of Neutral-Will Cases<sup>231</sup>:

*Not-Them Cases:* C is a not-them case if and only if the individual who brought about the behavior was not her/himself when s/he behaved.

*Not-Capable Cases:* C is a not-capable case if and only if the individual who

---

<sup>230</sup> My rendering pays close attention to Strawson (1962, 77 – 78). They are not in analysis form in the original, and their awkwardness results from my attempt to stay true to Strawson's wording. These are now more commonly put in terms of exemptions and excuses.

<sup>231</sup> See previous footnote. See, also, Strawson (1962, 78 – 79).

brought about the behavior was at the time mentally incapable of regarding her/himself or others in any moral sense.

Strawson is focused on the latter when he begins to discuss the objective stance in the lengthy passage quoted.

Rather than adopt OS-Pessimism, Strawson can distinguish between viewing someone as a natural, mechanistic thing and viewing someone as a *mere* natural, mechanistic thing. Those viewed as *mere* natural things could easily be taken to be individuals who show up in Strawson's Neutral-Will Cases. Strawson may have been on to this when, just before the lengthy passage quoted above from which the OS-Pessimist seems to draw their view, he writes:

The second and more important subgroup of cases allows that the circumstances were normal, but presents the agent as psychologically abnormal—or as morally undeveloped. The agent was himself; but he is warped or deranged, neurotic or just a child. When we see someone in such a light as this, all our reactive attitudes tend to be profoundly modified. I must deal here in crude dichotomies and ignore the ever-interesting and ever-illuminating varieties of case. What I want to contrast is the attitude (or range of attitudes) of involvement or participation in a human relationship, on the one hand, and what might be called the objective attitude (or range of attitudes) to another human being, on the other. Even in the same situation, I must add, they are not altogether *exclusive* of each other; but they are, profoundly, *opposed* to each other. (Strawson, 1962, 79, his emphasis)

The profound opposition may lie in the fact that we are at first blush curious whether and if so how encountering the world as full of only natural and mechanistic things preserves the various facets of peculiarity and wonder that we experience as we interact interpersonally and personally. Some may struggle to see how they can continue seeing

their brother's act as one of admiration or their partner's words as awash in love due to other attitudes that hold those individuals as somehow in part non-natural and non-mechanistic. But, upon reflection we may (and I think we should) diminish and silence that opposition as we see that such a perspective does nothing on its own to release us from the strains of commitment Strawson highlights. Indeed, we may come to see that the involvement and participation with each other that brought our reactive attitudes to life is not predicated on some metaphysical thesis about how much different we are from nature and mechanisms but on whether that nature and mechanism is too immature, broken, ill-wired, or out-of-order in some way.

Of course, whether Strawson saw that he didn't need OS-Pessimism is beside the point as is whether he succeeds. I dove into Strawson because this is the oft cited source of OS-Pessimism. Upon inspection we found that the no-difference thesis is *not* necessary for Strawson's argument. As long as he can mark out certain cases as the only cases in which suspension of the reactive attitudes makes sense—e.g., when interacting with a *mere* natural thing—, Strawson is poised to show that in all other cases our reactive attitudes are appropriate (perhaps, sometimes required) and can be so even in the face of determinism and on non-consequentialist grounds.

Thus, the OS-Pessimist cannot use this part of Strawson as a source of inspiration to motivate their position. Rather, Strawson *could have* appealed to OS-Optimism which firmly rejects the mistaken no-difference thesis. As I show in the next section, this repeats with respect to Strawson's argument for the practical inconceivability of completely ridding ourselves of reactive attitudes. The development of this optimistic replacement for the rejected pessimistic premise thought to be a key aspect of Strawson-

style accounts of moral responsibility and common ground between non-skeptics and skeptics will complete my argument. The result is an optimism that paves the way for a superior Strawsonian picture of holding and being morally responsible.

## 2: PARTICIPANT AND REACTIVE ATTITUDES

Rejecting OS-Pessimism might remove a central part of the *letter* of Strawson's argument in support of the

*Practical Inconceivability Thesis* (PIT): It would be, as a practical matter, inconceivable to completely jettison the reactive attitudes.

But, it doesn't touch the heart of the *spirit* of Strawson's intriguing thought. Rather, an OS-Optimist sympathetic to Strawson's vision simply needs an argument in favor of PIT that takes a path compatible with her optimism. In this section, I make perspicuous such a route, though it is latent in what I've already said.

Together with OS-Pessimism, Strawson seems to employ the following in support of PIT<sup>232</sup>:

- (1) The extent to which we adopt the objective stance determines the extent to which we suspend reactive attitudes.
- (2) We have an unshakable commitment to personal relationships.

Strawson's idea is fairly straightforward. Generally, Strawson is thinking that if we have

---

<sup>232</sup> My characterization of Strawson's argument for PIT strongly follows Shabo's (2012, 132). As he notes, it is a rough characterization, but adequate. Cf. Strawson (1962, esp. 79 – 83).

an unshakeable commitment to things that are themselves essentially tied to some other things, then we have an unshakable commitment to these latter things as well.<sup>233</sup> For him, the purported dissonance between personal relationships to which we have an unshakable commitment and the objective stance was meant to highlight the essential connection between personal relationships and the reactive attitudes. Traditionally, the disharmony is taken to stem from two things: (a) by adopting the objective stance an individual suspends the expectations and demands that prime the reactive attitudes, and (b) those expectations and demands and the attitudes they prime are central to deeply personal relationships. So, one might think that by rejecting (a) by adopting OS-Optimism we have rejected the core tie between something unshakable (i.e. personal relationships) and the reactive attitudes. One might conclude, then, that OS-Optimism runs contrary to PIT and, so, is antithetical to the Strawsonian picture.

The conclusion is mistaken. To see this, we should direct our attention to another theme that seems to run strongly through Strawson's essay and the views of Strawsonians generally. Strawsonians hold that being in a personal relationship with another involves seeing her as a co-participant in the moral community, that is, someone to be reasoned with rather than merely managed and controlled.<sup>234</sup> When we wonder why we should think that it is practically impossible to banish the reactive attitudes, Strawsonians point to two contrasting connections. One is the purported tight connection between lacking

---

<sup>233</sup> Again, nothing is lost if we put this in terms of desires, although it isn't clear that Strawson would recognize it immediately as his own if we did.

<sup>234</sup> Strawson (1962, 84 – 85). Cf. Stern (1974), Watson (1987, 229 – 230), Stephen Darwall (2006, Chpts. 4 and 5), McKenna (2005; 2012, 9 – 14 & 74 – 78), and Shoemaker (2007). I follow Macnamara (2011) in understanding this as a stance—namely, the participant stance.

susceptibility to the reactive attitudes and the objective stance.<sup>235</sup> And, the other is the tight connection alleged to hold between the reactive attitudes and the broad psychological orientation that we typically take toward ourselves and each other which colors us as beings whose behaviors can evince varying degrees of regard—i.e. the *participant stance*.<sup>236</sup> As we saw above, Strawson and Strawsonians focus on the objective stance and conclude that we cannot discard reactive attitudes anymore than we can exclusively take up the objective stance.

But, there is a superior option hiding in plain sight. As I already showed, exclusively adopting the objective stance leaves open whether that thing is *also* something simply to be managed and controlled rather than reasoned with. So, on the route I have in mind, (1) should be replaced with something like:

- (1\*) The extent to which we adopt the *participant stance* determines the extent to which we suspend reactive attitudes.

Alternatively, then, one focuses on the participant stance and concludes that we could not discard the reactive attitudes anymore than we could exclusively refrain from taking up the participant stance. This alternative foundation allows us to keep PIT for Strawsonian reasons. Furthermore, we keep those things without relying on the mistaken no-difference thesis or OS-Pessimism. On my Strawsonian picture we simply rely on what is apparently true; the optimistic view that the individual who adopts the objective stance may also adopt the participant stance and so care personally about the behaviors of others as I do care about my brother's and Gena's and they care about mine.

---

<sup>235</sup> Strawson (1962, 79 – 84).

<sup>236</sup> Strawson (1962, 79 – 84).

Strawsonians and non-Strawsonians, then, should happily embrace OS-Optimism. Still, given that OS-Pessimism is entrenched in the Strawsonian tradition, worries might persist. Although I doubt that I will tend to all of them, I conclude this section with a responsive eye to concerns that are salient to me. These concerns relate to Strawson's overall argument and certain implications of OS-Optimism. Attending to them will further strengthen my claim that OS-Optimism should replace OS-Pessimism in Strawson-style accounts of moral responsibility.

One might worry, for instance, that nixing the role played by the objective stance in a Strawsonian argument for PIT will render PIT impotent against Strawson's original targets. "Look," someone might say, "the objective stance is meant to make sense of the world envisioned by optimists and pessimists about the survival of our moral practices in the face of *determinism*. And, Strawson's main line against the heart of each view is the tight connection he thinks holds between the objective stance, personal relationships, and reactive attitudes. If you sever that connection, PIT can no longer be usefully employed against them."

In response, I concede that nothing I have said shows that Strawson's overall argument against such optimism and pessimism is *sound*.<sup>237</sup> Still, I don't see why severing the connection Strawson posits would effectively disassemble a Strawsonian argument against those views. The optimism Strawson sought to undermine simply claimed that to take others as causally determined would not undercut the fitness of our moral practices due to the fact that they would still be justified, say, as a deterrent. And, the pessimism Strawson sought to undermine was also focused on the upshot of our being

---

<sup>237</sup> For a nice collection of current studies of Strawson's essay, see, McKenna & Russell (2008).

causally determined. Strawson attempts to lend some credence to the idea that there are times when we are not emotionally involved with ourselves and others while maintaining that determinism does not serve as a standing defeater for such involvement. Thus, so long as something can play this sort of concessionary role, Strawson's argument remains. And, there is no reason to suppose that OS-Optimism rules out filling that position. Indeed, for example, a stance that paints others as *mere* natural, mechanistic things can easily serve this function and OS-Optimism leaves open that such a stance exists.

Following up, one might press that adopting the objective stance colors individuals as not *deserving of* our reactions to their behavior.<sup>238</sup> And, they might claim, a world void of such desert is a world without fitting reactive attitudes. Thus, they aver, the desert-tainted reactions I might have toward Hugh's or Gena's behavior are incoherent or inappropriate so long as I take the objective stance toward her. So, the thought continues, my OS-Optimism is untenable.

Consider, however, that this concern also relies on the mistaken assumption that where an individual is oriented toward everyone as natural and mechanistic she does not expect or demand anything of anyone such that the behavior of individuals she encounters will not fittingly trigger any desert-tainted response from her and so any such attitude or its expression will not be apt.<sup>239</sup> The mistake as we have seen is the failure to

---

<sup>238</sup> See, e.g., Sommers (2007, esp. 324). It is unclear whether Sommers discussing the desert of our *unexpressed* responsibility reactions, the *expression* of our reactions, or both. I think what I have to say covers each case. Cf. Pereboom (2014, esp. Chpts. 6 & 8)

<sup>239</sup>First, OS-Optimism is *silent* about whether the reactive attitudes are cognitive attitudes in the sense that they have some sort of propositional content. So, 'reasons' should be broadly construed. For non-cognitive conditions on the aptness of such attitudes, see, e.g., Gibbard (1990, Chpt. 7) and D'Arms & Jacobson (forthcoming). For cognitive conditions, see, e.g. Greenspan (1988), Wallace (1994, esp. Chpts. 5 & 6), and Nussbaum (2001, esp. Chpts. 1 & 2). Second, nothing about OS-Optimism restricts one from allowing that public expression of these attitudes is permissible where the permissibility conditions may

recognize that an individual who exclusively adopts the objective stance must still decide whether anyone is a participant in the moral community; she must decide whether, for instance, they are to be reasoned with or merely ushered about. Hence, adopting the objective stance does not itself rule out being deserving or worthy of, say, praise or blame.<sup>240</sup>

Finally, an alternative objection stems from a specific view about the reactive attitudes. For instance, R. Jay Wallace claims that narrowing the class of reactive attitudes to those attitudes of resentment, indignation, and guilt “clarifies and sharpens the opposition between objectivity of stance and the reactive emotions” (1994, 32).<sup>241</sup> Wallace’s thought echoes his distinction between attitudes within the broad class of emotions to which one is typically susceptible in personal relationships.<sup>242</sup> On his account, the *reactive* emotions are just those that are grounded by an expectation that one behave in a particular manner, namely, in a way that evinces a certain sort of regard for herself or others. And, ostensibly, Wallace thinks that one does not have such expectations of those who she views from the objective stance.<sup>243</sup> One might object, then, that OS-Optimism forces us to collapse Wallace’s important distinction.

Notice that this objection simply assumes OS-Pessimism. On OS-Optimism, as I showed above I can have the sort of expectation of my brother or Gena that Wallace is concerned whether I encounter them always and everywhere as natural and mechanistic

---

come apart from the attitude’s fitness conditions. On these points in particular, see, e.g., T. M. Scanlon (2008, 166 – 179) and Angela M. Smith (2007).

<sup>240</sup> Though saying more is beyond the scope of this paper, this appears to be compatible with accepting what is typically referred to as a *basic desert thesis*. For starters, see McKenna (2012, Chpts. 5 – 7), Pereboom (2013; 2014, esp. Chpt. 6), and Boxer (2014, Chpts. 4 & 5) on desert. [omitted].

<sup>241</sup> Cf. Wallace (2011). [omitted]

<sup>242</sup> Wallace (1994, 30 – 33).

<sup>243</sup> Wallace (1994, 31 n. 17)..

and so respond with resentment or indignation only where such expectations are in play. Thus, restricting the reactive emotions to Wallace's three is not antithetical to my project. Rather, his restriction might usefully help us clarify and sharpen the opposition between responding to an individual as responsible and being oriented toward her as something to be managed rather than reasoned with.

Therefore, despite these interesting concerns, OS-Optimism remains and finds a home in Strawsonian thought. Indeed, the idea that the extent to which an individual responds to herself and others as morally responsible is not dependent on whether she experiences herself or them as natural mechanisms but on whether she experiences herself or them as a mere natural mechanism—i.e., something that is not a participant in the moral community or something to be managed and controlled as opposed to reasoned with, respected, or loved—further secures that Strawsonian home against skepticism in at least two ways. It gets rid of a mistaken keystone thesis and it offers an alternative view that serves to mend gaps left by rejection of that thesis. Hence, OS-Optimism and views about participation in the moral community already accepted by Strawsonians work in unison to fashion an improved foundation for Strawson-style accounts.

### 3: CONCLUSION

As we have seen, according to the picture delivered by OS-Optimism, whether an individual responds to herself or others as morally responsible is not dependent on whether she experiences herself or them as natural or mechanistic but, rather, on whether she experiences herself or them as, e.g., things that are not participants in the moral

community or things to be managed and controlled as opposed to reasoned with, deeply cared for, respected, or loved. Though an individual may encounter herself and others as natural and mechanistic, she can respond to and value the attitudes expressed by her behavior and theirs in ways that are appropriate or fitting. Thus, as is natural, she may come to love or befriend someone as she navigates her vulnerability to feelings like resentment, indignation, guilt, gratitude, admiration, or pride. And, contrary to tradition, this love and friendship can have grounds beyond the superficiality of what she can get out of it or how these relationships benefit her or some group. Hence, OS-Optimism finds a home in Strawsonian thought and further secures that home by ridding it of a mistaken keystone thesis and thereby canceling an important concession traditional Strawsonians make on which encountering a world full of natural mechanisms leaves us cold and bereft of the passion required for deep, interpersonal relationships. Hence, OS-Optimism and views about participation in the moral community already accepted by Strawsonians work in unison to fashion an improved foundation for Strawson-style accounts of moral responsibility. Though it remains to be seen whether those accounts will outlast all criticism, with OS-Optimism they are more formidable if for no other reason than they allow what is true—namely, that we can love each other and find depth in our friendships whether oriented toward ourselves and others as natural mechanisms or not. In the next chapter, I bring this and the other aspects of the recognition account forward to construct a Strawsonian argument against global skepticism.

## VIII: CONCLUSION: SKEPTICISM AND RESENTMENT

For better or worse, responsibility responses are central to our moral experience. The controversial question is not *whether* we respond to individuals as morally responsible but whether anyone actually *is* morally responsible. The chapters in this volume are stand-alone essays insofar as none relies on the main thesis of another for support. However, they paint a non-skeptical picture of moral responsibility and provide important insights for arguing against global skepticism about moral responsibility. In this concluding essay, I base an argument against global skepticism on that non-skeptical picture. Roughly, my main thesis here is that, if I am correct, then the proponent of global skepticism cannot fall back on the truth of the claim that no one is ultimately responsible.

More precisely, my goal is to do what many think cannot be done. I support a rejection of global skepticism by employing a style of argument most widely attributed to P. F. Strawson famous work, “Freedom and Resentment”. To see how I aim to argue, consider first the general argument issued in support of global skepticism:

*The Skeptical Argument:*

- P1) Responding to an individual S in way W as an agent responsible for  $\phi$  is appropriate only if S deserves (or is fit) to be responded to in way W for  $\phi$ -ing.
- P2) No one deserves (or is fit) to be responded to as a responsible agent (i.e. No one deserves (or is the fit subject of) responsibility responses).  
(Anti-Fitness Thesis)

- C1) Therefore, responding to anyone as a responsible agent is always inappropriate.
- P3) If responsibility responses are always inappropriate, then no one is responsible.
- C2) Therefore, no one is responsible.

Typically the main point of conversation is (P2), i.e. the Anti-Fitness Thesis. Proponents of the conclusion that no one is responsible lean heavily on the idea that in order to deserve or be fit to be the subject of a responsibility response for  $\phi$ -ing an individual must be ultimately responsible for  $\phi$ . They go on to argue that no one is ultimately responsible for anything. Unfortunately these proponents of global skepticism are not univocal about the conditions for being ultimately responsible for something. However, they all seem to endorse the idea that being ultimately responsible for something means or entails that the agent was *the source* of that thing or that its causal history terminates in a *wholly self-made agent*. And, they all endorse the view that causal determinism rules out sourcehood or being a wholly self-made agent.

I focus on (P2) as well and eventually get around to undercutting the appeal to ultimate responsibility. My argument against that popular appeal is Strawsonian in style. Thus, if I succeed I will have shown two things. First, I will have shown that (P2) is not supported by the appeal to ultimate responsibility. Second, my argument will demonstrate that a Strawson-style argument can be successfully employed against global skepticism.

More perspicuously, in “Freedom and Resentment”, Strawson attempts to reconcile (a) the pessimistic view that our moral responsibility concepts and practices are

incoherent and unjustified in the face of causal determinism with (b) the optimistic view that we retain a sort of consequentialist coherency and justification of them despite the truth of causal determinism. The reconciliation is meant to show that our moral responsibility practices and concepts survive the truth of causal determinism without the need to interpret them simply in terms of mechanisms employed only to control and manage each other toward a common good. One main line against Strawson criticizes his reluctance to really raise and answer the

*External Question:* Whether we are psychologically incapable of ridding ourselves of the practice of holding each other and ourselves responsible or not, is anyone the fit subject of moral-responsibility responses, (i.e. is the practice of responding to individuals as normatively responsible for what they do externally justified)?<sup>244</sup>

Strawson comes close to raising the question but ultimately and to the dissatisfaction of many has this to say about it:

...if we could imagine what we cannot have, *viz.* a choice in this matter, then 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. (Strawson, 1962, 83, his emphasis)

R. Jay Wallace succinctly characterizes the dissatisfaction many have with Strawson's position stating that he wrongly assumes that "the only possible reasons that might be cited for or against our practice of holding people morally responsible are pragmatic

---

<sup>244</sup> See, e.g., Wallace (1994, 95 – 109) and essays in McKenna & Russell (2008).

reasons, concerning the gains and losses of the practice for human life” (Wallace, 1994, 100).

Generally, the worry for Strawson-style arguments is that they only offer an *internal* justification of responsibility responses like praise and blame rather than an *external* justification. And, the thought goes on; we need the latter in order to reject anything like the Anti-Fitness Thesis (i.e., (P2)). What is meant by ‘internal’ and ‘external’ varies somewhat depending on who is criticizing the Strawson-style argument. I argue that none of the current forms of this general worry are sound (§1). My basic point throughout is that those rejecting the Strawson-style arguments fail to appreciate the resources available to proponents of that style of argument. I follow up by analyzing K. E. Boxer’s (2013) Strawson-style argument. I argue that Boxer’s reconstruction is insightful, but ultimately does not go through (§2). I go on, then, to offer my own reconstruction (§3).

The reconstruction I put forward leans on three things. First, I nuance the Strawsonian-picture so that it is clear that we can ask whether determinism undermines the appropriateness of our responsibility practices—i.e., we can ask the External Question. Second, I argue that, since (a) the attitudes involved in responsibility responses track manifestations of will and (b) we have no good reason to believe that determinism renders it impossible for individuals to manifest wills of good or ill quality, the answer to the External Question is negative—causal determinism does not undermine the appropriateness of responsibility responses. Finally, I appeal to Elizabeth Lane Beardsley’s (1960) key but overlooked idea that asking whether an individual is *ultimately* responsible for  $\phi$  only makes sense if we can sensibly raise and affirmatively

answer questions about her status as an individual normatively responsible for her acts and attitudes in other normative realms of interpersonal life. The upshot is that the question of ultimate responsibility is sensible as is its affirmative answer only if our responsibility responses where ultimate responsibility is not at issue are appropriate. Thus, unlike other reconstructions of the Strawson-style argument, we preserve the significance of questioning ultimate responsibility. I turn, then, to investigate two of the most well-pronounced rejections of the Strawson's original argument.

### 1: STRAWSON'S TWO PART ARGUMENT

Generally, the route authors take toward rejecting Strawson involves peeling apart two strands in his argument and rejecting each separately. Call this, the *Divide-and-Conquer Strategy*. Most identify a line of Strawson's that they term the *Naturalistic Argument*. Roughly, the Naturalistic Argument concludes that, as a practical matter, it would be inconceivable to go in for a wholesale rejection of the reactive attitudes on the basis of claims about our psychological and moral-social nature. But, authors interpret Strawson differently when it comes to the second main line of argument. For instance, in their favored terminology, Paul Russell distinguishes between the Naturalistic Argument and the *Rationalistic Argument*. Alternatively, R. Jay Wallace identifies what he terms the *Pragmatic Argument* as Strawson's second main line. In my critique, then, I will canvass objections to the Naturalistic Argument and those two varieties of the second main line of thought.

The basic complaint against the Naturalistic Argument is that it wrongly assumes that we cannot really coherently ask the External Question. And, even though the other line of argument whether construed *ala* Russell (i.e. the Rationalistic Argument) or Wallace (i.e. the Pragmatic Argument) in some way raises or attempts to raise the External Question, many claim that the answer Strawson offers is unsatisfying.

In this section, I argue that this Divide-and-Conquer Strategy only weakens Strawson's argument by simply misrepresenting it. Once we allow ourselves to put the main lines of argument together, we will see where Strawson's argument fails and be in a position to put forward a powerful reconstruction of that argument.

### 1.1: The Naturalistic Argument

Consider the somewhat obscure Naturalistic Argument attributed to Strawson. The basic argument runs something like this:

- N1) The practice of holding individuals responsible (i.e. reacting to the behavior of ourselves and others with the reactive attitudes) requires justification, only if it is at least practically possible to go in for a wholesale rejection of holding responsible (i.e. reacting to the behavior of ourselves and others with the reactive attitudes). (*Naturalist Premise*)
- N2) It is, as a practical matter, not possible to go in for a wholesale rejection of holding responsible (i.e. it is practically inconceivable to go in for such a rejection). (*Practical Inconceivability Thesis*)
- NC1) Therefore, the practice of holding individuals responsible (i.e. reacting to the behavior of ourselves and others with the reactive attitudes) requires no justification. (*Naturalist Conclusion*)

Both premises are called into question even by those who basically adopt a Strawsonian framework in the end.<sup>245</sup>

The main complaint, however, focuses on the Naturalist Premise. Paul Russell, in particular, objects that Strawson's Naturalistic Argument of argument trades on an ambiguity in both the pessimistic view Strawson ultimately seeks to reject as well as Strawson's own naturalism as expressed in (N2), i.e. the Practical Inconceivability Thesis. Recall that the pessimist holds that our practices of holding responsible are incoherent and unjustified in the face of determinism. According to Russell, we can and should distinguish between *type*-pessimism and *token*-pessimism as well as *type*-naturalism and *token*-naturalism about the reactive attitudes. Russell uses fear as an example to make the distinction clear. I lay out the example in Table 1:

	PESSIMISM		NATURALISM	
	Type	Token	Type	Token
Fear	Given that there is no external/rational justification for the fact that we are susceptible or liable to fear, we (a) <i>can</i> and (b) <i>must</i> free ourselves of this (irrational) disposition to fear.	Given our circumstances, we are never justified in being afraid (i.e. we are never justified in entertaining any tokens of fear)	Our liability to fear is natural to humans and the fact that we are susceptible or liable to fear requires no justification.	We regularly, naturally, and inevitably encounter circumstances in which fear is entirely appropriate or reasonable.

**Table 1**<sup>246</sup>

<sup>245</sup> See, e.g., Wallace (1994, 31 & 38 – 40).

<sup>246</sup> Russel (1992).

To generate versions of type and token pessimism as well as naturalism relevant to our discussion, we simply replace fear with the reactive sentiments.

Russell's idea can be best understood by clearly marking the difference between the two sorts of pessimisms he delineates and the two sorts of naturalism. Consider first the difference between Russell's pessimisms. To see the difference in each brand of Russell's pessimisms, we need to follow the

*Basic Ingredient:*     The claim that there is no external/rational justification for x.

The difference, then, between type-pessimism and token-pessimism is that the former takes the basic ingredient *to imply something* about us whereas the latter takes the basic ingredient *to follow from something* about us. Each claims that something about our responses to x of the Basic Ingredient requires no justification. Type-naturalism claims that our liability to experience x requires *no justification* while token-naturalism simply claims that our experiences of x are appropriate.

Russell argues that type-naturalism does not even engage the pessimist and token-naturalism fairs no better. According to Russell, token-naturalism is what Strawson must maintain to establish the conclusion of his Naturalistic Argument via the Naturalist Premise and the Practical Inconceivability Thesis. But, Russell maintains that, if Strawson "is embracing token-naturalism, then, worse still, he is embracing a position that is committed to suspect and disturbing factual claims and which, moreover, does not even address itself to the (legitimate) concerns of the Pessimist" (Russell, 1992, 151). Expanding, Russell writes:

What is particularly disturbing about Strawson's naturalistic strategy [Natural Argument], expressed in more general terms, is that it casts doubt on our ability, or capacity to curb or control our emotional life according to the dictates of reason. More specifically, it seems clear that, despite disclaimers to the contrary, Strawson's naturalistic strategy invites us to accept or reconcile ourselves to reactive attitudes (and their associated retributive practices) even in circumstances when we have to reason to repudiate them. (Russell, 1992, 152)

Russell concludes, then, that despite Strawson's claims to the contrary "[we] have every reason to take the Pessimist seriously, and this puts greater weight on Strawson's rationalistic strategy [Rationalistic Argument]" (Russell, 1992, 152). In other words, the only way to go is to answer the External Question head on.<sup>247</sup>

Russell's careful analysis notwithstanding, he gives us no reason to be suspicious of Strawson's strategy. The main problem with Russell's argument is the thing that is *generally* wrong with the ubiquitous Divide-and-Conquer Strategy.<sup>248</sup> To see this, note that we can remove the 'disturbing results' by putting the pieces of Strawson's argument back together. Strawson certainly thinks that doing without the practice of holding responsible is, as a practical matter, inconceivable. He also thinks that type-naturalism is true (though not under that description, perhaps). However, Strawson also seems to have a positive answer to the question, "Are we ever justified in having the attitudes constitutive of holding an individual responsible?"<sup>249</sup>

---

<sup>247</sup> Cf. Wallace (1994, 96 – 103).

<sup>248</sup> The divide-and-conquer strategy is used by some to put Strawson's view aside completely. It is used by others to motivate development of it.

<sup>249</sup> My reconstruction, then, will not run the course of Magil's (1997) defense. Roughly, Magil holds that those opposed to Strawson wrongly assume that we should focus on justifying the attitudes and opts for focusing on when the attitudes are appropriate. Whether we cast the debate in terms of justification or appropriateness we can press the same issues that Russell and others press. So, Magil will need my reconstruction as well.

Strawson's positive answer to that question is that the attitudes constitutive of holding responsible are the reactive attitudes. These attitudes track the quality of the will an individual manifests through action. Strawson goes on to offer an account of when individuals are properly excused or exempt from these attitudes. That is, he goes on to offer an account of the reasons for and against having a reactive attitude toward someone. Hence, on the one hand, Strawson offers a story on which we are liable to experience the reactive attitudes due to how much we care about the interpersonal attitudes that others hold regarding us as persons. That is his type-naturalism. It may not offer a justification for our liability to the reactive attitudes, but it certainly offers an explanation of their presence.

On the other hand, Strawson offers a story about the propriety of the reactive attitudes, which contra Russell engages the pessimist and is not committed to disturbing factual claims. On Strawson's view, manifestations of will are the reasons for which we have certain reactive attitudes. Or, in more naturalistic terms, they trigger our liability to react with attitudes like resentment and indignation. Here, the quality of an individual's will is determined by the extent to which her decisions are sensitive to reasons of mutual respect. And, a reactive attitude is fitting only where its content is properly sensitive to the quality of will manifested—i.e. that which the reactive attitudes track. Finally, as Strawson sees it, determinism neither excuses nor exempts. He concludes, then, that our moral responsibility concepts and practices, which centrally involve the reactive attitudes and their expression, are fitting even in the face of causal determinism. Given that (a) the care we have for each as persons grounds our liability to the reactive attitudes and (b) we experience those attitudes appropriately only in circumstances where that care was either

properly taken into consideration or not, Strawson does not commit us to anything disturbing. Rather, we should hope that something like Strawson's view is correct, since it would be disturbing to find out that our reactive attitudes (which can sometimes be harsh especially when expressed) are randomly triggered by things like one's race or religion or, worse still, if we found out that they were fitting when so triggered.

Of course, perhaps the Russell is correct that the pessimist will remain unsatisfied to some extent. For one, they may not accept a fitting-attitude story in place a justification story. But, it's unlikely that all pessimists are also against fitting-attitude explanations, especially where that explanation aligns with the view that we should not be treating people differently on the basis of things like their race, religion, height, etc. Of course, the pessimist may have further (legitimate) concerns. For instance, the pessimist might still want Strawson to support what he seems to assume—namely, that the care that under girds our liability to the reactive attitudes is appropriate or justified in the face of causal determinism. At most, then, Russell has pointed out that Strawson's argument seems incomplete. But, that incompleteness does not support the view that we should abandon this sort of argument. It simply calls out for us to complete it. I turn, then, to further draw out the incompleteness of Strawson's original argument by addressing the concerns raised against Strawson's other line of argument.

## 1.2: The Second Line

Recall that authors diverge over what they see as the other main line of argument in Strawson. I will focus on two versions. Discussion of each will highlight important

problems that should not show up in a viable reconstruction of Strawson's argument. Also, the discussion reveals elements that I will carry forward to my own reconstruction.

### 1.2.1: *The Rationalistic Version*

Russell claims that Strawson's other main line of argument involves identifying excusing and exempting conditions in a manner that is meant to show that determinism does not bear on the appropriateness of the reactive attitudes. Russell takes Strawson to be distinguishing between *specific* conditions that would render the reactive attitudes inappropriate and *global* conditions. According to Russell, the specific conditions do not call into question the agent's status as a fit subject of the reactive attitudes but the global conditions do. And, for Russell, Strawson holds that an individual who meets the global conditions that render the reactive attitudes inappropriate must be psychologically abnormal in such a way that she is not part of the community of morally responsible agents. But, as Russell points out, Strawson claims that this is not the *normal* case. For Strawson, typically people can and do deliberate in a manner that is more or less sensitive to reasons of mutual respect and regard. Finally, on Russell's interpretation, Strawson holds that determinism is incoherent if it entails that the normal condition is that we are *all* psychologically abnormal. The conclusion of this Rationalistic Argument that Russell attributes to Strawson, then, is that the truth of determinism does not call into question fitness for the reactive attitudes and, so, does not call into question their propriety.

Russell argues that this Rationalistic Argument collapses. To see this, Russell thinks we need to substitute certain words for others in Strawson's argument. According

to Russell, replacing “the abnormal” and “abnormality” with “the incapacitated” and “incapacity” does two things.<sup>250</sup> First, according to Russell, the replacement more accurately captures the pessimist’s concern. On his view, the pessimists are concerned that determinism undermines the moral capacity required to be fit objects of our reactive attitudes. Russell understands the pessimist to hold, for instance, that the “relevant capacity [...] is “free will” or “contra-causal freedom”” (Russell, 1992, 154). Second, the replacement makes salient the incompleteness of Strawson’s argument. Russell maintains that Strawson’s observation that libertarian accounts of freedom involve “panicky metaphysics” may “succeed in casting doubt on one interpretation of what the relevant capacities are supposed to be [...but] [it] is far from obvious [... that] that in itself establishes that the truth of the thesis of determinism poses no threat to our moral capacities and hence to our reactive attitudes” (Russell, 1992, 154). He concludes, then, that Strawson’s Rationalistic Argument “is, as it stands, at best incomplete” (Russell, 1992, 154).

Russell and I agree that Strawson’s argument is incomplete. However, Russell’s concern boils down to the claim that Strawson’s argument is incomplete because Strawson does not show that determinism is compatible with freedom. But, that either sets the bar too high or simply misunderstands the debate.<sup>251</sup>

The pessimist may well think that free will and causal determinism are incompatible. Nevertheless, contra Russell, Strawson is simply addressing the pessimistic claim that the concepts and practices involved in moral responsibility responses are incoherent and unjustified where determinism obtains. Since Strawson

---

<sup>250</sup> Russell (1992, 154).

<sup>251</sup> Cf. Boxer (2013, 82 n. 60).

thinks that holding responsible is a matter of tracking and reacting to the regard manifested in our decisions, he needs to show that determinism gives us no reason to think that the causally determined cannot manifest regard. Strawson certainly seems to think that this is the case. However, his argument is incomplete because he offers no reason why we should agree. If this does not lower the bar, it at least gets us talking about the right bar.

### 1.2.2: *The Pragmatic Version*

Strawson's overall argument is incomplete in another way as well. To see this, consider the argument that Wallace attributes to Strawson, i.e. the Pragmatic Argument.<sup>252</sup> As Wallace points out, Strawson *does* raise the External Question even after concluding that jettisoning our practice of holding responsible is practically inconceivable. Not only does Strawson raise the question of whether we should continue to hold individuals responsible, he also directs us to the considerations to which we must attend in order to answer the question:

...if we could imagine what we cannot have, *viz.* a choice in this matter, then 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. (Strawson, 1962, 83, his emphasis)

Wallace's objection is that Strawson simply and wrongly assumes that "the only possible reasons that might be cited for or against our practice of holding people morally

---

<sup>252</sup> Wallace (1994, 99).

responsible are pragmatic reasons, concerning the gains and losses of the practice for human life” (Wallace, 1994, 100).

Wallace makes a good point here and one that further establishes that Strawson’s argument is incomplete. But, like Russell, Wallace presses the complaint that Strawson’s argument is incomplete as an argument meant to settle the debate between compatibilists and incompatibilists about moral responsibility.<sup>253</sup> That is simply not what Strawson is up to.

As an argument that attempts to reconcile the pessimist and the optimist, Strawson’s argument is incomplete in at least three key ways. First, as the discussion of Russell brought out, we need to support the view that we have no good reason to believe that determinism undermines one’s ability to decide on the basis of reasons of mutual respect and regard. In other words, the argument lacks support for the view that individual’s can evince certain qualities of will despite determinism. Second, Wallace is correct—we need an argument supporting the assumption that the only way to determine whether we should abandon our practice of holding responsible is to appeal to the gains or losses such abandonment would bring to our human lives. Finally, we need something to bridge the gap between (a) the conclusion that the truth of determinism does not bear on the appropriateness of our practice of holding responsible and (b) the notion that the only version of the External Question left to answer is one where the answer will only involve the upshot of pragmatic considerations. My reconstruction will attempt to meet all three points thereby offering a complete version of Strawson’s argument and one that

---

<sup>253</sup> Wallace (1994, Chpt. 4, esp. 103 – 117).

finds room for the worry of ultimate responsibility which is at the heart of libertarian concerns. I start to rebuild with Elizabeth Lane Beardsley's help.

## 2: Beardsley's Argument

Beardsley's (1960) treatment of the pessimist is slightly and importantly different from Strawson's. On her view, the pessimists are basically correct but their position incomplete.<sup>254</sup> Beardsley's idea is that in asking about the propriety of holding responsible we must distinguish three different genuinely *moral* perspectives. In two of the three perspectives we are concerned with determining whether their behavior or character rises to/fails to meet a certain normative standard.<sup>255</sup> Beardsley sees the answers to questions about an individual's praise-/blameworthiness and questions about the excellence of one's character as answers that correctly take into account only the *local* causal circumstances.<sup>256</sup> However, the third perspective involves taking into account the *global* causal circumstances.<sup>257</sup> Beardsley claims that the answer from this third perspective is the same for everyone: the ultimate cause of actions and the ultimate cause of coming to possess a particular character is one over which no one has voluntary control.<sup>258</sup> Thus, our question, "Is anyone morally responsible for their act/character?" from this third perspective really is the question, "Is anyone *ultimately* morally responsible for their act/character?", and is answered negatively for everyone. Hence, it

---

<sup>254</sup> Beardsley (1960, 4, 11, 14 – 15).

<sup>255</sup> Ibid., 4 – 11.

<sup>256</sup> Ibid., 12 – 15.

<sup>257</sup> Ibid.

<sup>258</sup> Ibid.

is from this third perspective that pessimists ask their questions and answer them pessimistically.

However, according to Beardsley, the answer that S is not ultimately morally responsible for  $\phi$  is legitimate just in case  $\phi$  is an act for which S possesses positive/negative moral worth or moral credit.<sup>259</sup> That is, the claim that S is not ultimately morally responsible for  $\phi$  makes sense just in case a statement about S is true from one of the other two perspectives. On Beardsley's view, then, the question of whether one is ultimately responsible for a particular piece of behavior or a particular quality of her character is incoherent when we cannot attribute either to her as a basis of moral appraisal in these other local perspectives.

I want to put Beardsley's simple but seemingly true point to work to fill a lacuna in Strawson's argument. So doing will help us find room for a Beardsley sort of take on optimism as well. Thus, it will thereby help us do what Strawson seems not to have done—namely, find a place for the significance of ultimate responsibility.

More precisely, then, I want to use Beardsley's point to move from

I: The thought (developed below) that determinism does not undermine our ability to manifest regard, and, so, does not undermine attributing actions to agents as the basis of moral appraisal;

to

II: The thought that the only external question left to answer is one whose answer will depend on the gains and losses to human life.

---

<sup>259</sup> Ibid., 14 – 15 & 15 n. 21.

In a nutshell, my thought is that after settling (I), Beardsley's position shows that we have no other question to ask about whether the attitudes involved in moral responsibility responses are ever justified. So, only after settling (I) we can legitimately ask other sorts of questions about our moral responsibility concepts and practices like those about being ultimately responsible. Those questions are significant for understanding our place in the universe. However, they do not bear on the coherency and justifiability of our moral responsibility practices. Rather, that coherency and justification is what makes raising the question about ultimate responsibility sensible.

Further, then, according to (II) the optimist Strawson was concerned with is primarily operating at the same global perspective from which we ask ultimate responsibility questions. Since the answer to the questions about ultimate responsibility is likely negative, we might preserve the concepts and practice from this perspective only by pointing to the value of the distribution of consequences where we do so preserve them. But, that does not change the concepts into consequentialist concepts nor does it render the propriety of our practices consequentialist in nature. It is simply a way to determine whether essentially non-consequentialist concepts and practices should continue. My critique of K. E. Boxer's reconstruction of the Strawson-style argument will help illuminate these thoughts and further progress toward my own reconstruction.

### 3: RECONSTRUCTIONS

The main focus is to develop a Strawson-style argument that meets the global skeptic head on by using the insights gathered from the foregoing as well as those gathered from

a recent attempt to do a similar rebuild. Here are the three insights from above that a Strawson-style argument must incorporate in order to directly address the concerns of the skeptic:

- 1: It must support the view that we have no good reason to believe that determinism or lack of ultimate responsibility undermines one's ability to decide on the basis of reasons of mutual respect and regard;
- 2: It must support the assumption that the only way to determine whether we should abandon the concepts and practices centrally involved in moral responsibility responses is to appeal to the gains or losses such abandonment would bring to our human lives; and
- 3: It must contain something plausible to bridge the gap between (a) the conclusion that the truth of determinism does not bear on the coherency and appropriateness of our moral responsibility concepts and practices and (b) the notion that the only version of the External Question left to answer is one where the propriety it concerns is (at in large part) answered by considering the gains and losses of continued deployment of moral responsibility concepts and practices.

To make this more concrete, I first consider Boxer's reconstruction. I argue that she fails on each point. But, seeing how she misses the mark is important primarily because Boxer's rebuild is basically ingenious and comes very close to doing (1) – (3).

### 3.1: K. E. BOXER'S RECONSTRUCTION

Boxer's project is very much like my own insofar as we are both interested in moving forward against The Skeptical Argument from a Strawson-style argument. In her attempt

to do so, Boxer focuses on undermining the following main supporting for the No Desert

Thesis:

Ultimate Responsibility Argument:

UR1) An individual, S, deserves (or is fit) to be responded to in way W for  $\phi$ -ing only if s/he is ultimately responsible for  $\phi$ . (*Ultimate Responsibility Thesis*)

UR2) No one is ultimately responsible for anything. (*No Ultimate Responsibility Thesis*)

URC1) Therefore, no one deserves (or is fit) to be responded to in way W for  $\phi$ -ing.

Authors support the Ultimate Responsibility Thesis and the No Ultimate Responsibility Thesis in various ways. And, some adopt the former but not the latter.<sup>260</sup> Boxer rejects each.

I will not rehearse the entirety of Boxer's insightful discussion. Instead I want to focus on her defense of Strawson's argument. Consider, then, Boxer's overarching argument:

B1) The Ultimate Responsibility Thesis is true only if being ultimately responsible for  $\phi$  is a necessary condition of deserving to be held liability-responsible.<sup>261</sup>

B2) Being ultimately responsible for  $\phi$  is not a necessary condition of deserving to be held liability-responsible.<sup>262</sup>

B3) Therefore, the Ultimate Responsibility Thesis is false.

---

<sup>260</sup> See, e.g., Robert Kane (1996).

<sup>261</sup> Boxer (2013, 3 – 7, 34 – 35, & 138 – 144).

<sup>262</sup> Boxer (2013, 8 – 9, 60 – 61, 89 – 90, & 138 – 144).

In support of (B2), Boxer addresses two general theories about holding responsible. One of those theories is Strawson's

*Reactive Attitudes Account:* A moral responsibility response is constituted by a complex conative-affective attitude (e.g. reactive attitudes like resentment, indignation, guilt, gratitude, approbation, and pride).<sup>263</sup>

Boxer argues that Strawson's Reactive Attitudes Account does not require that one is ultimately responsible in order to be appropriately held responsible. So, if this is the correct way to understand being held liability-responsible for  $\phi$ , then it is not the case that being ultimately responsible for  $\phi$  is a necessary condition for being appropriately so held.<sup>264</sup>

More precisely, Boxer argues that if ultimate responsibility is a necessary condition for being the appropriate object of the (moral) reactive attitudes, then an individual's being ultimately responsible is one of the *internal conditions* that must be met in order for her/him to be the appropriate object of the reactive attitudes.<sup>265</sup> Boxer explains that Strawson requires that one be "a normal adult human being capable of participating in normal adult interpersonal relationships and of understanding the [demand that those with whom we (or others) are interpersonally related treat us (and each other) with a certain degree of regard and respect] that arises from, and regulates, just these relationships" (Boxer, 2013, 80, *footnote omitted*). In conclusion, she writes:

---

<sup>263</sup> Boxer (2013, Chpt. 3)

<sup>264</sup> Boxer (2013, 62).

<sup>265</sup> Boxer (2013, Chpt. 3, esp. 65 – 66).

That the capacity to engage in ordinary adult interpersonal relationships as we normally understand them is the capacity crucial to an agent's being an appropriate target of the moral reactive attitudes and the demand underlying them means that we can be confident that no metaphysical thesis 'consistent with the facts as we know them' would render the moral reactive attitudes universally inappropriate by their own internal standards. This includes any thesis concerning an absence of ultimate responsibility that might hold true of us as we are. Among the facts as know them is that many of us are capable of participating in ordinary adult interpersonal relationships as we normally understand them. For many of us do participate in such relationships. [...] It is this [...] that explains Strawson's confidence that the thesis of determinism poses no threat to desert of the moral reactive attitudes. We can be equally confident that no thesis concerning the absence of ultimate responsibility that might be true of us as we are would pose such a threat. (Boxer, 2013, 82, *footnote omitted*)

So, according to Boxer, being ultimately responsible is not one of the internal conditions that must be met on Strawson's view. Thus, the proponent of the Ultimate Responsibility Thesis would gain nothing by adopting Strawson's reactive attitudes account as the correct account of holding one liability-responsible.

My primary concern with Boxer's support of (part of) her premise (B2)—i.e. the claim that being ultimately responsible for  $\phi$  is not a necessary condition of deserving to be held liability-responsible—is that it wrongly assumes that once we identify the internal mechanisms that produce responsibility responses like the reactive attitudes or once we cite what it is that these attitudes are responsive to where no reference to ultimate responsibility is made in that identification our job is done.<sup>266</sup>

To see why that assumption is incorrect consider a parody. Suppose we find out that beliefs about a colored object are produced by or respond to appearances. Surely, it

---

<sup>266</sup> I say 'part of' because, as I said, Boxer also addresses another account of holding liability-responsible.

is illicit to move from that fact to the fact that no further question needs to be raised about the veracity of our beliefs about colored objects. For one, the appearance might be due to colored lights shining on the object and making it appear to the believer to be a color that it isn't. Likewise, suppose Boxer is correct that our reactive attitudes are produced by or respond to what appear to be individuals exercising their capacity to interact with us or with each other. We cannot move from that fact to the view that no further question needs to be raised about the fitness of our attitudes. Here, though they look to be exercising a capacity to interact, perhaps they aren't. Or, more plausibly, perhaps an individual has that capacity but cannot exercise it in a manner that makes those seemingly interpersonal activities attributable as the basis of moral appraisal. This is where the theorist with sympathies to ultimate responsibility sneaks back into the picture. Like the epistemologist who is careful to point out that having a reason to believe that *p* is insufficient for dubbing the belief true, those who favor the view that ultimate responsibility is required for proper attribution will claim that having a reason to respond to individuals as morally responsible for what they've done is insufficient for dubbing that response fitting.<sup>267</sup>

---

<sup>267</sup> For instance, they might think that a capacity that is required for normal adult interaction is the capacity to be the source of your actions, where having such a capacity requires ultimate responsibility. Or, more plausibly perhaps, they could think that beyond the capacity of normal adult human interaction another condition internal to determining the appropriateness of the reactive attitudes is that one must be the source of her actions, again, where being such requires being ultimately responsible. Alternatively, they could attempt to cash this capacity out in terms of possessing alternative possibilities. But, Boxer (2013, 2) thinks of the source condition as more basic and more plausibly the right condition for proponents of the Ultimate Responsibility Thesis to fight for. In any case, notice, on her way to establishing (B1)—i.e. the claim that the Ultimate Responsibility Thesis is true only if being ultimately responsible for  $\phi$  is a necessary condition of deserving to be held liability-responsible—Boxer does *not* reject sourcehood as a reasonable understanding of what it takes to either flourish in normal adult interaction or be the appropriate target of the reactive attitudes. Nor does Boxer show that one can be the source of one's actions absent

Thus, Boxer's defense is inadequate. First, it fails to meet adequacy condition (1)—i.e., Boxer does not show that we have no good reason to believe that determinism or lack of ultimate responsibility undermines the fitness of the reactive attitudes. Second, Boxer's view leaves untouched the idea that the only external question left to ask is one that is answered by the gains and losses doing so has for human flourishing—i.e., she fails to meet adequacy condition (2). Finally, then, Boxer's reconstruction fails to meet adequacy condition (3) insofar as it does not bridge the gap between (a) the conclusion that the truth of determinism does not bear on the coherency and appropriateness of our moral responsibility concepts and practices and (b) the notion that the only version of the External Question left to answer is one where the propriety it concerns is (at least in large part) answered by considering the gains and losses of continued deployment of moral responsibility concepts and practices. Thus, the fact that Boxer's account comports well with my analysis of Russell and Wallace notwithstanding, we must look elsewhere for an adequate reconstruction.

### 3.2: MY RECONSTRUCTION

The pieces for my reconstruction are basically already on the table. The key pieces to it are (a) establishing that determinism alone gives us no reason to believe that our practice of responding to individuals as responsible is unjustified, and (b) securing the view that,

---

ultimate responsibility. Rather, Boxer simply argues that proponents of the Ultimate Responsibility Thesis need a new argument to support that thesis regardless of how they cash out what it is to be ultimately responsible. Thus, nothing Boxer says here rules out the move to adding this sort of nuance to the fitness conditions of the reactive attitudes.

once (a) is established, we have no reason to believe that the answer to the only external question left to ask turns on whether we have ultimate responsibility.

The basic strategy that moral responsibility skeptics use to support the Anti-Fitness Thesis (i.e. No one deserves (or is the fit subject of) responsibility responses) is to show that no account adequately shows that we have good reason to believe that anyone deserves (or is the fit subject of) responsibility responses. Typically, the strategy is employed against a type of view. The idea, then, is that if successful their argument impugns all accounts of that type. The skeptic repeats until all reasonable account-types have been addressed and shown to fail. Hence, to push against a Strawson-style account, the skeptic would need to pick on a quality of will view of moral responsibility. I start my reconstruction, then, by noting the ways in which the contributions in this dissertation defend against that sort of push. These contributions help highlight the Strawson-style path to rejecting the Anti-Fitness Thesis in a manner that meets adequacy conditions (1) – (3).

Roughly, quality of will accounts of moral responsibility hold that *S* is morally responsible for  $\phi$  if and only if  $\phi$  manifests who *S* is as a practical agent—i.e. manifests *S*'s will. Further, *S* is morally praiseworthy/blameworthy for  $\phi$  if and only if  $\phi$  manifests a good/ill quality of will. There are at least two common ways to pick on such an account. One way is to call into question the adequacy of the manifests relation to do the work it is assigned to do. Another way is to show that the moral appraisals grounded by the view are lacking somehow. The idea on this latter route is that one must do more than just show that we are morally responsible in this quality of will sense in order to overturn the Anti-Fitness Thesis in virtue of the fact that proponents of that thesis are not

concerned to show that no one deserves *superficial* forms of moral appraisal. Rather, they are concerned to show that no one deserves or is the fit subject of *non-superficial* or deep forms of moral appraisal.

In Chapter 6, I dealt with the objection that negligent acts reveal the general inadequacy of the manifests relation and so shows the failure of quality of will views to account for the moral responsibility and praise-/blameworthiness of individuals. I argued for a causal interpretation of the relation according to which the quality of that manifested will is partly determined by what is absent in it. Thus, manifesting *ill* will does not require the individual to have any negative attitude toward the individual injured by her action or individuals in general. I rejected the view that lacking an ability to understand the interpersonal exchange of which responsibility responses are a piece was fully excusing or exculpatory. Thus, where one has what it takes to manifest a will of good/*ill* quality she is in the community of morally responsible agents and so, all things equal, her manifestations of will are fit for moral responsibility responses.

In Chapter 5, I addressed the objection that the most prominent type of quality of will view, namely, Self-Revelation Views are inadequate due to the fact that the moral appraisals they warrant are superficial. I appealed to the idea that depth of an appraisal of S for  $\phi$  is determined by whether it was warranted where such an appraisal is warranted only where  $\phi$  is the result of, is sustained by, or is otherwise directly under the control of S's agential activity. Since an individual must have the capacity to render, sustain, or otherwise have direct control over  $\phi$  in order to have attitudes and acts attributed to her as the basis of moral appraisal and so manifest a will with morally evaluable qualities, the moral appraisals warranted on my Self-Revelation View which is a significant type of

quality of will view are deep not superficial. Furthermore, this view identifies the capacity I think must be exercisable if an individual is to be fit for moral responsibility responses.

In Chapter 7 I argue for a new foundation for Strawson-style accounts of moral responsibility. According to the new foundation, we should be optimistic about the survival of our deep interpersonal relationships in the face of universally adopting the broad psychological stance that paints everyone as natural and mechanistic. I called that view Objective-Stance Optimism. Traditionally, Strawsonians and most skeptics hold that such adoption undermines those relationships because it silences fitting morally responsibility responses. By rejecting Objective-Stance Pessimism, then, I severed the common ground between Strawsonians and those skeptics.

But trouble loomed. My rejection of Objective-Stance Pessimism seemed to put Strawson's Naturalistic Argument in jeopardy, since many take it to rely on the view that universal adoption of the objective stance undercuts the propriety of the reactive attitudes. However, I saved that argument by showing that contrary to popular belief universal adoption of the objective stance still allows us to encounter individuals as participants in the moral community. There I relied on the Strawsonian theme that such interpersonal engagement is really what primes our liability to respond to each other and ourselves as morally responsible, since it is that engagement that shows that we understand each other and ourselves as individuals to be reasoned with rather than merely pushed around and controlled.

More than that, along the way to reestablishing the crucial link between responsibility responses and certain broad psychological stances I showed that we have

no reason to believe that encountering everyone everywhere always as natural and mechanistic destroys the basis for viewing ourselves as beings responsive to reasons of mutual regard and respect. It is that argument that helps us meet the first condition of adequacy—namely, showing that we have no reason to believe that determinism or lack of ultimate responsibility renders us unfit for the reactive attitudes.

Very important for my argument here is Nomy Arpaly's observation that it would be surprising if the efficacy of content depended on the truth of causal determinism. I used Arpaly's insight for at least two different purposes in the argument of Chapter 7. First, I used it to shut down an alternative avenue to rejecting the idea that we retain the grounds for fitting responsibility responses where we exclusively take up the objective stance. Second, I used it to further support Objective-Stance Optimism by calling into question the supposition that there is no difference between seeing something as a natural, mechanistic thing and seeing it as a *mere* natural, mechanistic thing. The upshot is that we need not worry that a wider community of morally responsible agents pushes us to include things that we see as natural, mechanistic things. We can expand the community in this way and still offer the Strawson-inspired reconciliation.

I want to reuse Arpaly's point here to support the view that we have no good reason to think that our practice of responding to individuals as responsible for what they do should be given up in the face of our lack of ultimate responsibility. Given the observation that it would be utterly surprising to discover that the efficacy of content depends on the truth of causal determinism, causal determinism should give us no reason to think that actions are not attributable to agents as the basis of moral appraisal. The reason is that all we need for such attribution is for actions to issue from one's agential

faculties in a manner that allows us to say truly of them that they were caused by the content of the reasons to which the agent was sensitive to in coming to her decision to act. Thus, it seems even in the presence of causal determinism we have no reason to suspect that individuals cannot manifest wills of good or poor quality, since this simply requires us to respond to the content of reasons as one does when he cries at the touch of a sad song.<sup>268</sup>

Our lack of ultimate responsibility should not make us no more suspicious of our agential capacities than causal determinism does. Whether ultimately responsible or not we will encounter and recognize reasons for and reasons against and we will respond to them in a reliable pattern according to our normative orientation—i.e., aligned with our beliefs, values, desires, cares, preferences, etc.<sup>269</sup> True, we may not be gods in the sense that we shape ourselves from conception or birth into the beings we are today. But, that is not the significance of ultimate responsibility. If it were, it would also be comically the insignificance of ultimate responsibility as well. What would be significant to moral responsibility is the failure to have the capacity to encounter and recognize reasons that push us to change components of our normative orientation our responsiveness allows us the leeway to reinvent ourselves as we go along. But, that failure is abnormal. And, if the thesis of causal determinism or the main line of proponents of ultimate responsibility is that such a failure is normal and universal, then both are muddled and incoherent at their core.

---

<sup>268</sup> I do not mean for this to exclude the requirement that one have the capacity of normative competence. Rather, I take it that this sort of responsiveness requires normative competence.

<sup>269</sup> Fischer's (2006) terminology.

Thus, we should not give up our moral responsibility practices in the face of causal determinism or lack of ultimate responsibility. But, we did not reach this conclusion by reflecting on whether it would be impossible to give up the practice. Rather, causal determinism (or naturalism) or lack of ultimate responsibility simply gives us no reason to think that the practice is suspect. The only thing that should lead us to the Anti-Fitness Thesis is reason to believe that individuals are incapable of manifesting wills of good or poor quality.

Finally, we complete this rejection of the Anti-Fitness Thesis with Beardsley's insight. The External Question remains, i.e., from the perspective of ultimate responsibility, is the practice of responding to agents as morally responsible inappropriate or unjustified and if so should we give it up if we can? Given the foregoing, we can shrug with Beardsley as we point out that we couldn't sensibly take up that perspective were the practice inappropriate or unjustified. Further, as Strawson suggests the answer to the question of ultimate responsibility lies not in deep and mystical metaphysics about libertarian freedom and causal determinism. We've already done the metaphysical work necessary for establishing the coherency and justification of our moral responsibility concepts and the local attitudes and practices in which they feature without dressing them in consequentialist or non-naturalistic clothes. That result came from thinking about the nature of praise and blame as well as the nature of agency and what threatens it. From this global perspective we simply have to determine what it is that is all things considered best for our flourishing. Our answer there will be our answer to the External Question. The significance of ultimate responsibility is that in thinking about it we are forced to consider things and ourselves as we truly are—*viz.*, accidents waiting to happen—while

acknowledging that we eventually do happen in deliberate ways and it is those happenings that are attributable to us as the basis of the sorts of deep moral appraisals central to our responsibility responses and interpersonal lives.

## REFERENCES

- Alvarez, Maria. 2010. *Kinds of Reasons: An Essay in the Philosophy of Action*. Oxford: Oxford University Press.
- Anscombe, G. E. M. 1957. *Intention*. Oxford: Blackwell.
- Arneson, Richard. 'The Smart Theory of Moral Responsibility and Desert'. in Olsaretti, Serena (ed.), *Desert and Justice*. Oxford: Oxford University Press: 233 – 263.
- Arpaly, Nomy. 2003. *Unprincipled Virtue: An Enquiry into Moral Agency*. Oxford: Oxford University Press.
- \_\_\_\_\_. 2006. *Merit, Meaning, and Human Bondage: An Essay on Free Will*. New Jersey: Princeton University Press.
- Arpaly, Nomy & Schroeder, Timothy. 1999. 'Praise, Blame, and the Whole Self'. *Philosophical Studies*, 93: 161 – 188.
- Baker, Brenda M. 1987. 'Mens Rea, Negligence, and Criminal Law Reform'. *Law and Philosophy*, 6: 53 – 88.
- Barry, Melissa. 2005. 'Humean Theories of Motivation'. in Shafer-Landau, R. (ed.), *Oxford Studies in Metaethics*, v. 5: Oxford: Oxford University Press: 195 – 224.
- Beardsley, Elizabeth Lane. 1960. 'Determinism and Moral Perspectives'. *Philosophy and Phenomenological Research*, 21: 1 – 20.
- \_\_\_\_\_. 1960. 'Determinism and Moral Perspectives'. *Philosophy and Phenomenological Research*, 21/1: 1 – 20.
- \_\_\_\_\_. 1978. 'Blaming'. *Philosophia*, 8/4: 573 – 583.
- Bedke, Matthew. Unpublished. 'A Case for Besires'. Faculty.arts.ubc.ca/mbedke/Research.html: accessed, 1.18.14.
- Bennett, Jonathan. 1980. 'Accountability'. in Straaten, Z. (ed.), *Philosophical Subjects*. Oxford: Oxford University Press: 14 – 47.
- Benson, Paul. 1987. 'Freedom and Value'. *Journal of Philosophy*, 84: 465 – 486.
- Boxer, K. E. 2013. *Rethinking Responsibility*. Oxford: Oxford University Press.
- Brady, James B. 1980. 'Recklessness, Negligence, Indifference, and Awareness'. *Modern Law Review*, 43: 381 – 399.
- \_\_\_\_\_. 1996. 'Conscious Negligence'. *American Philosophical Quarterly*, 33: 325 – 335.
- Coleman, Mary Clayton. 2008. 'Directions of Fit and the Humean Theory of Motivation'. *Australasian Journal of Philosophy*, 86/1: 127 – 138.
- Collins, John. 1988. 'Belief, Desire, and Revision'. *Mind*, 97/387: 333 – 342.

- D'Arms, Justin & Jacobson, Daniel. (forthcoming). *Rational Sentimentalism*. Oxford: Oxford University Press.
- Darwall, Stephen. 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, MA: Harvard University Press.
- de Tarde, Gabriel. 1912. 'The Theory of Moral Responsibility'. in Morris, H. (ed.), *Freedom and Responsibility: Readings in Philosophy and Law*. California: Stanford University Press: 46 – 49.
- Deigh, John. 2011. 'Reactive Attitudes Revisited'. in Bagnoli, C. (ed.), *Morality and the Emotions*. Oxford: Oxford University Press: 197 – 216.
- Dennet, Daniel C. 1984. *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge, MA: MIT Press.
- Downie, R. S. 1966. 'Objective and Reactive Attitudes'. *Analysis*, 27: 33 – 39.
- Ebels-Duggan, Kyla. 2013. 'Dealing with the Past: Responsibility and Personal History'. *Philosophical Studies*, 164: 141 – 161.
- Edgerton, Henry W. 1926. 'Negligence, Inadvertence, and Indifference: The Relation of Mental States to Negligence'. *Harvard Law Review*, 39: 849 – 870.
- Ekstrom, Laura Waddell. 2000. *Free Will: A Philosophical Study*. Boulder, CO: Westview Press.
- Feinberg, Joel. 1969. 'Sua Culpa'. Feinberg, J. (ed.), *Doing and Deserving*, 1970. New Jersey: Princeton University Press: 187 – 221.
- \_\_\_\_\_. 1970. *Doing and Deserving*. Princeton: Princeton University Press.
- Fischer, Agneta H. & Manstead, Antony S. R. 2008. 'Social Functions of Emotions'. in Lewis, M. Haviland-Jones, J., & Feldman Berrett, L. (eds.), *The Handbook of Emotions*, 3<sup>rd</sup> edition. New York: The Guilford Press: 456 – 468.
- Fischer, John Martin. 2006. 'The Cards That You Are Dealt'. *Journal of Ethics*, 10: 107 – 129.
- Fischer, John Martin & Ravizza, Mark. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
- Frijda, Nico H. 2008. 'The Psychologists' Point of View'. in Lewis, M. Haviland-Jones, J., & Feldman Berrett, L. (eds.), *The Handbook of Emotions*, 3<sup>rd</sup> edition. New York: The Guilford Press: (79)
- Gibbard, Alan. 1990. *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press.
- Glover, Jonathon. 1970. *Responsibility*. Oxford: Oxford University Press.
- Greenspan, Patricia S. 1988. *Emotions and Reasons: An Inquiry into Emotional Justification*. New York: Routledge.
- \_\_\_\_\_. 2001. 'Genes, Electrotransmitters, and Free Will'. in Wasserman, D. and Wachbroit (eds.), *Genetics and Criminal Behavior*. New York: Cambridge University Press: 243 – 258.
- Griffiths, Paul. 2004. 'Is Emotion a Natural Kind?'. in Solomon, R. C. (ed.), *Thinking About Feeling: Contemporary Philosophers on Emotions*. Oxford: Oxford University Press: 233 – 249.

- Haji, Ishtiyaque. 1998. *Moral Appraisability: Puzzles, Proposals, and Perplexities*. Oxford: Oxford University Press.
- Hájek, Alan & Pettit, Philip. 2004. 'Desire Beyond Belief'. *Australasian Journal of Philosophy*, 82/1: 77 – 92.
- Hart, H. L. A. 1968. 'Negligence, *Mens Rea*, and Criminal Responsibility'. in Hart, H. L. A. (ed.), *Punishment and Responsibility*. Oxford: Oxford University Press: chpt. 6.
- Heller, Kevin Jon and Dubber, Markus D. 2011. *The Handbook of Comparative Criminal Law*. Stanford, CA: Stanford University Press.
- Helm, Bennet W. 2011. 'Responsibility and Dignity: Strawsonian Themes'. in Bagnoli, C. (ed.), *Morality and the Emotions*. Oxford: Oxford University Press: 217 – 234.
- Hieronymi, Pamela. 2004. 'The Force and Fairness of Blame'. *Philosophical Perspectives*, 18: 115 – 148.
- \_\_\_\_\_. 2011. 'Making a Difference'. *Social Theory and Practice*, 37: 81 – 94.
- Hume, David. 1777. *An Enquiry Concerning the Principles of Morals*. LaSalle, IL: Open Court Publishing (1991).
- Hurley, Elisa & Macnamara, Coleen. 2010. 'Beyond Belief: Toward a Theory of the Reactive Attitudes'. *Philosophical Papers*, 39: 373 – 399.
- Husak, Doug. 2011. 'Negligence, Belief, Blame, and Criminal Liability: The Special Case of Forgetting'. *Criminal Law and Philosophy*, 5: 199 – 218.
- King, Matt. 2009. 'The Problem with Negligence'. *Social Theory and Practice*, 35: 577 – 595.
- Korsgaard, Christine M. 1996. *Creating a Kingdom of Ends*. Cambridge: Cambridge University Press.
- Lazarus, Richard. 1991. *Emotion and Adaptation*. Oxford: Oxford University Press.
- Levy, Neil. 2005. 'The Good, the Bad, and the Blameworthy'. *Journal of Ethics and Social Philosophy*, 1: 2 – 15.
- Lewis, David. 1988. 'Desire as Belief'. *Mind*, 97/387: 323 – 332.
- \_\_\_\_\_. 1996. 'Desire as Belief II'. *Mind*, 105/418: 303 – 313.
- Lunney, Mark and Oliphant, Ken. *Tort Law: Text and Materials*. Oxford: Oxford University Press.
- Macnamara, Coleen. 2011. 'Holding Others Responsible'. *Philosophical Studies*, 152: 81 – 102.
- \_\_\_\_\_. 2013a. 'Skrew you!' and 'Thank you!'. *Philosophical Studies*, 165: 893 – 914.
- \_\_\_\_\_. 2013b. 'Taking Demands Out of Blame'. Coates, D. J. and Tognazzini, N. (eds.), *Blame: Its Nature and Norms*. Oxford: Oxford University Press: 141 – 161.
- \_\_\_\_\_. *forthcoming*. 'Reactive Attitudes as Communicative Entities'. *Philosophy and Phenomenological Research*.

- Magill, Kevin. 1997. 'Blaming, Understanding, and Justification: A Defense of Strawson's Naturalism about Moral Responsibility'. in McKenna, Michael and Russell, Paul (eds.), *Free Will and Reactive Attitudes: Perspectives on P. F. Strawson's Freedom and Resentment*, 2008. Burlington, VT: Ashgate Publishing: 187 – 200.
- McGeer, Victoria. 2013. 'Civilizing Blame'. in Tognazzini, N. and Coates, D. Justin (eds.), *Blame: Its Nature and Norms*. Oxford: Oxford University Press: 162 – 189.
- McKenna, Michael. 2005. 'Where Frankfurt and Strawson Meet'. *Midwest Studies in Philosophy*, 29: 163 – 180.
- \_\_\_\_\_. 2012. *Conversation and Responsibility*. Oxford: Oxford University Press.
- McKenna, M. and Russell, P. 2008. *Free Will and Reactive Attitudes: Perspectives on P. F. Strawson's Freedom and Resentment*. Burlington, VT: Ashgate Publishing.
- Midgley, Mary. 1984. *Wickedness*. New York, NY: Routledge.
- \_\_\_\_\_. 1994. *The Ethical Primate*. New York, NY: Routledge.
- Nichols, Shaun. 2007. 'After Incompatibilism: A Naturalistic Defense of the Reactive Attitudes'. *Philosophical Perspectives*, 21: 405 – 428.
- Nussbaum, Martha. 2001. *Upheavals of Thought*. Cambridge: Cambridge University Press.
- Oshana, Marian A. L. 1997. 'Ascriptions of Responsibility'. *American Philosophical Quarterly*, 34: 71 – 83.
- \_\_\_\_\_. 2004. 'Moral Accountability'. *Philosophical Topics*, 32: 255 – 274.
- Pereboom, Derk. 2001. *Living without Free Will*. Cambridge: Cambridge University Press.
- \_\_\_\_\_. 2009. 'Free Will, Love, and Anger'. *Ideas y Valores: Revista de Colombiana de Filosofía*, 141: 5 – 25.
- \_\_\_\_\_. 2013. 'Free Will Skepticism, Blame and Obligation'. in Tognazzini, N. and Coates, D. J. (eds.), *Blame: Its Nature and Norms*. Oxford: Oxford University Press: 189 – 206.
- \_\_\_\_\_. 2014. *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press.
- Pettit, Philip. 2001. *A Theory of Freedom: From the Psychology to the Politics of Agency*. Cambridge: Polity Press.
- Price, Huw. 1989. 'Defending Desire-as-Belief'. *Mind*, 98/389: 119 – 127.
- Prinz, Jesse. 2004. *Gut Reactions: A Perceptual Theory of Emotion*. Oxford: Oxford University Press.
- Russell, Paul. 1992. 'Strawson's Way of Naturalizing Responsibility'. in McKenna, M. and Russell, P. (eds.), *Free Will and Reactive Attitudes: Perspectives on P. F. Strawson's Freedom and Resentment*, 2008. Burlington, VT: Ashgate Publishing: 143 – 157.
- \_\_\_\_\_. 2004. 'Responsibility and the Condition of Moral Sense'. *Philosophical Topics*, 32: 287 – 306.

- Salovey, Peter, Detweiler-Bedell, Brian, Detweiler-Bedell, Jerusha B., & Mayer, John D. 2011. 'Emotional Intelligence'. in Lewis, M. Haviland-Jones, J., & Feldman Berrett, L. (eds.), *The Handbook of Emotions*, 3<sup>rd</sup> edition. New York: The Guilford Press: 533 – 547.
- Scanlon, T. M. 1988. 'The Significance of Choice'. in McMurrin, S. (Ed.), *The Tanner Lectures on Human Values*, v. 8. Salt Lake City: University of Utah Press: 149 – 216.
- \_\_\_\_\_. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- \_\_\_\_\_. 2008. *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge, MA: Harvard University Press.
- Shafer-Landau, Russ. 2003. *Moral Realism: A Defense*. Oxford: Clarendon Press.
- Sher, George. 2006. *In Praise of Blame*. Oxford: Oxford University Press.
- \_\_\_\_\_. 2009. *Who Knew?: Responsibility without Awareness*. Oxford: Oxford University Press.
- Shabo, Seth. 2012. 'Incompatibilism and Personal Relationships: Another Look at Strawson's Objective Attitude'. *Australasian Journal of Philosophy*, 90: 131 – 147.
- Shoemaker, David. 2007. 'Moral Address, Moral Responsibility, and the Boundaries of the Moral Community'. *Ethics*, 118: 70 – 108.
- \_\_\_\_\_. 2011. 'Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility'. *Ethics*, 121: 603 – 632.
- \_\_\_\_\_. Forthcoming a. 'Qualities of Will'. *Social Philosophy & Policy*, 30/1 – 2: 95 – 120.
- \_\_\_\_\_. Forthcoming b. *Responsibility at the Margins*. Oxford: Oxford University Press.
- Shoemaker, David, and Tognazzini, Neal. 2015. *Oxford Studies in Agency and Responsibility, Vol. 2: 'Freedom and Resentment at 50'*. Oxford: Oxford University Press.
- Smart, J. J. C. 1961. 'Free Will, Praise, and Blame'. in Watson, Gary (ed.), *Free Will*, 2<sup>nd</sup> edition, (2003). Oxford: Oxford University Press: 58 – 71.
- Smilansky, Saul. 2000. *Free Will and Illusion*. Oxford: Oxford University Press.
- Smith, Angela M. 2000. 'Identification and Responsibility'. in van den Beld, A. (Ed.), *Moral Responsibility and Ontology*. Dordrecht: Kluwer Academic Publications: 233 – 246.
- \_\_\_\_\_. 2005. 'Responsibility for Attitudes: Activity and Passivity in Mental Life'. *Ethics*, 115: 236 – 271.
- \_\_\_\_\_. 2007. 'On Being Responsible and Holding Responsible'. *Journal of Ethics*, 11: 465 – 484.
- \_\_\_\_\_. 2008. 'Control, Responsibility, and Moral Assessment'. *Philosophical Studies*, 138: 367 – 392.
- \_\_\_\_\_. 2012. 'Attributability, Answerability, and Accountability: In Defense of a Unified Account'. *Ethics*, 122: 575 – 589.
- Smith, Craig A. & Lazarus, Robert S. 1991. 'Appraisal Components, Core Relational Themes, and the Emotions'. *Cognition and Emotion*, 7: 233 – 269.
- Smith, Holy. 1983. 'Culpable Ignorance'. *Philosophical Review*, 92: 543 – 571.

- \_\_\_\_\_. 2011. 'Non-Tracing Cases of Culpable Ignorance'. *Criminal Law and Philosophy*, online first.
- Smith, Michael. 1987. 'The Humean Theory of Motivation'. *Mind*, 96/381: 36 – 61.
- \_\_\_\_\_. 1988. 'On Humeans, Anti-Humeans, and Motivation: A Reply to Pettit', *Mind*, 97/388: 589 – 595.
- \_\_\_\_\_. 1994. *The Moral Problem*. Oxford: Blackwell Publishing.
- Sobel, David & Copp, David. 2001. 'Against Direction of Fit Accounts of Belief and Desire'. *Analysis*, 61/1: 44 – 53.
- Sommers, Tamler. 2007. 'The Objective Attitude'. *The Philosophical Quarterly*, 57: 321 – 341.
- Stern, Lawrence. 1974. 'Freedom, Blame, and Moral Community'. *The Journal of Philosophy*, 71: 72 – 84.
- Stocker, Michael. 1979. 'Desiring the Bad: An Essay in Moral Psychology'. *Journal of Philosophy*, 76/12: 738 – 753.
- Strawson, P. F. 1962. 'Freedom and Resentment'. in Watson, Gary (ed.), *Free Will*, 2<sup>nd</sup> edition, (2003). Oxford: Oxford University Press: 72 – 93.
- Sverdlik, Steven. 1993. 'Pure Negligence'. *American Philosophical Quarterly*, 30: 137 – 149.
- Swartzler, Steven. 2013. 'Appetitive Desires and the Fuss about Fit'. *Philosophical Studies*, 165: 975 – 988.
- \_\_\_\_\_. Unpublished. 'A Challenge for Humean Internalism'. Under review.
- Taylor, Paul C. 2004. *Race: A Philosophical Introduction*. Cambridge: Polity Press.
- Terry, Henry T. 1915. 'Negligence'. *Harvard Law Review*, 29: 40 – 54.
- Vargas, Manuel. 2005. 'The Trouble with Tracing'. *Midwest Studies in Philosophy*, 29: 269 – 291.
- \_\_\_\_\_. 2013. *Building Better Beings*. Oxford: Oxford University Press.
- Walker, Margaret Urban. 2003. *Moral Contexts*. Lanham, MD: Rowman & Littlefield.
- Wallace, R. Jay. 1994. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press.
- \_\_\_\_\_. 2004. 'Normativity and the Will'. Wallace, R. J. (ed.), *Normativity and the Will*, 2006. Oxford: Oxford University Press: 71 – 81.
- \_\_\_\_\_. 2011. 'Dispassionate Opprobrium: On Blame and the Reactive Sentiments'. in Wallace, R. J., Kumar, R., and Freeman, S. (eds.), *Reasons and Recognition: Essays on the Philosophy of T. M. Scanlon*. Oxford: Oxford University Press 348 – 372.
- Watson, Gary. 1975. 'Free Agency'. in Watson, G. (Ed.), *Agency and Answerability: Selected Essays*, 2004. Oxford: Oxford University Press: 13 – 32.
- Watson, Gary. 1987. 'Moral Responsibility and the Limits of Evil'. in Watson, G. (ed.), *Agency and Answerability: Selected Essays*, (2004). Oxford: Oxford University Press: 219 – 259.
- \_\_\_\_\_. 1996. 'Two Faces of Responsibility'. in Watson, G. (Ed.), *Agency and Answerability: Selected Essays*, 2004. Oxford: Oxford University Press: 260 – 288.
- Weintraub, Ruth. 2007. 'Desire as Belief, Lewis Notwithstanding'. *Analysis*, 67/2: 116 – 122.

Wertheimer, Roger. 1998. 'Constraining Condemning'. *Ethics*, 108: 489 – 501.

Wolf, Susan. 1981. 'The Importance of Free Will'. *Mind*, 90: 386 – 405.

\_\_\_\_\_. 1987. 'Sanity and the Metaphysics of Responsibility'. in Schoeman, F. D. (ed.), *Responsibility, Character, and the Emotions*. Cambridge: Cambridge University Press: 46 – 62.

\_\_\_\_\_. 1990. *Freedom within Reason*. Oxford: Oxford University Press.

\_\_\_\_\_. 2011. 'Blame, Italian Style'. in Wallace, R. J., Kumar, R., and Freeman, S. (eds.), *Reasons and Recognition: Essays on the Philosophy of T. M. Scanlon*. Oxford: Oxford University Press: 332 – 347.

Zimmerman, Michael J. 1988. *An Essay on Moral Responsibility*. New Jersey: Rowman and Littlefield.

\_\_\_\_\_. 1997. 'Moral Responsibility and Ignorance'. *Ethics*, 107: 410 – 426.