

## Responsibility and the Limits of Good and Evil<sup>1</sup>

Robert H. Wallace, University of Arizona

P.F. Strawson's compatibilism about free will and his attendant account of moral responsibility have had a considerable influence on contemporary work.<sup>2</sup> This is so even among those who do not count themselves as followers of Strawson's overall compatibilist program. Even opponents address Strawson's main themes: the moral emotions and our practices of interpersonal expectations and demands.<sup>3</sup> Despite this influence, there are two damning problems for Strawson's theory. First, it appears to restrict the class of morally responsible agents to those who are in fact members of the moral community. However, as Gary Watson has argued in "Responsibility and the Limits of Evil" (2008/1987), this has a disturbing consequence: extreme evil *exempts* an agent from moral responsibility. If so, why shouldn't we view this as a *reductio*? Second, in some cases our emotional reaction to an evildoer's history clashes with our emotional expressions of blame. We might then worry that anyone's actions can be explained by his or her history, and thereby, conflict with our present blame. We might also worry that had our history been like the unfortunate evildoer, we too might have been evil. This compromises our standing to blame. Together, this would undermine the expression of blame, generally, and so be self-defeating. It is critical to the Strawsonian project to respond to these problems. I will do so in a novel way. As I see it, we have an impoverished vocabulary of reactive attitudes. Expanding our vocabulary can dissolve the problems. To do so, I will

---

<sup>1</sup> I would like to thank Mark Timmons, Jeremy Reid, and David Shoemaker for comments and feedback. I would especially like to thank Michael McKenna for his generous assistance. Finally, thanks to an anonymous reviewer at *Philosophical Studies* for thoughtful and constructive suggestions for improving the article.

<sup>2</sup> For example, see Bennett (1980), Darwall (2006), Fischer and Ravizza (1998), McKenna (1998, 2012), Russell (1992), (2004), Shoemaker (2015), Wallace (1994), Watson (1987), and Wolf (1981).

<sup>3</sup> For example, see Pereboom (2001), (2014) Smilansky (2000), and G. Strawson (1986).

revisit the example of extreme evil discussed by Watson, which motivates the challenges, and then provide a contrasting example of my own, one of extreme good. By examining these cases at two different ends of a spectrum, I will identify a peculiar pair of reactive attitudes. Relying upon them, I will then provide an unorthodox defense of Strawsonian compatibilism.

### 1. The Basics of Strawsonianism

Strawson's theory is about moral responsibility in the accountability sense of being on the hook for one's actions, so to speak.<sup>4</sup> It has three components. First, Strawson thinks that the practices of holding someone morally responsible are expressed by and consist of emotional *reactive attitudes*. Some emotions—resentment and indignation—are responsive to the intentions of other persons in a way that many emotions are not. We might get upset if our car is a lemon, but we couldn't be genuinely indignant *at* the car! We get indignant at, say, a co-worker when we find out that he or she has unfairly spoken ill of our other colleagues.

When holding another responsible by way of a reactive attitude like resentment, we are reacting to whether or not someone has shown us due consideration in their intended actions. This is the second component: moral responsibility tracks the *quality of will*—the good or ill will—with which a person acts. “Quality of will” is best understood as referring to the concern or lack thereof that a person shows to others, especially as manifested in her actions. This concern is constituted by an agent's attitudes and intentions (Strawson 2008/1962; McKenna 2012; 59-60; Shoemaker, 2015). In turn, in being disposed to the pertinent reactive attitudes, we *interpret* agents' actions as expressing these quality-of-will constituting attitudes and intentions.

---

<sup>4</sup> This kind of moral responsibility is opposed to moral responsibility in other senses, like moral assessment of character (attributability responsibility) or moral assessment of judgement-sensitive attitudes (answerability responsibility). For an extended discussion see Shoemaker (2015).

Why focus on interpretation? Sometimes we interpret each other's actions incorrectly. Hence, we have need for excuses, justifications, and exemptions. We could be *excused* from being held responsible—"I was shoved and fell into you!" Alternatively, we could provide *justification* for our actions—"if I didn't shove you that car would have hit you!" In either case, we are invited to see a different quality of will behind the (putative) injury. Exemptions show that someone is not the proper target of the reactive attitudes at all. An exempted agent is "*incapacitated* in some or all respects for ordinary inter-personal relations" (Strawson 2008/1962; 25). The exempted agent is seen through what Strawson calls objective attitudes. He or she has become an object that demands explanation or requires management, even if temporarily. This person is not expected to meet the demands of morality. Examples might include very young children or the seriously mentally ill.

Here is the third component of the view: there is nothing more to being morally responsible than being the appropriate target of reactive attitudes as set by our practices of holding one another responsible. Our expression of reactive attitudes signifies the moral demands we place on one another. To lack either the capacity for relevant quality of will or the capacity to engage in interpersonal relations would *dispel* the reactive emotions that express our moral demands and would invite an objective attitude.

We needn't go into detail about how these components make for compatibilism. Instead, our focus is the capacities needed to be a target of reactive attitudes. This is essential to responsible agents on the Strawsonian view; however, he does not specify what these capacities are. Building on a proposal first suggested by Gary Watson (1987), contemporary Strawsonians have attempted to advance his view by arguing that reactive attitudes communicate responses to interpretations of quality of will. This provides a way to precisify the requisite capacities. Call the following the argument from communication, which is due to

the interesting work of Colleen Macnamara: (1) morally responsible agents are eligible candidates for praise and blame. (2) Praise and blame in the form of the reactive attitudes are forms of communication and, in paradigmatic cases, take a praised or blamed agent as an intended addressee. (3) In order to be a candidate for praise and blame, then, one has to be a candidate addressee of praise or blame. Therefore, (4) in order to be a proper candidate for address, one must have the capacity to understand the meaning communicated by the reactive attitudes. What indicates that someone genuinely understands the meaning of a received reactive attitude? Typically, they need an understanding of what an appropriate emotional response to a given reactive attitude is. For instance, someone who knows that blame should be accompanied by guilt and a desire to make amends when faced with indignation is a proper target of address in the form of blame. They need not actually feel guilt.<sup>5</sup>

McKenna's (2012) view of the reactive attitudes shows us that persons with these capacities to understand reactive attitudes constitute a moral community (but importantly McKenna's view is that one could have these capacities *and* be outside the moral community—more on this later). By analogy to the linguistic competence needed to hold an intelligible conversation within a linguistic community, the morally responsible agent is involved in an interpretive enterprise whereby different patterns of conduct come to indicate states of mind, intentions, and so on. An agent embedded in these practices understands *both* their own actions and the actions of others by reference to the set of patterns that have come to indicate good or ill will. The knowledge of one requires the knowledge of the other. Why? In order for the reactive attitudes to work they need to be understood by members of the moral community

---

<sup>5</sup> This is an expanded paraphrase of Macnamara (2015)'s excellent characterization of the argument. I have made an amendment. Macnamara thinks that to be the appropriate target of address is to undergo the right sort of response to a reactive attitude. I disagree. People often understand what they should feel but do not concurrently have those feelings. Watson (2008/1987), Shoemaker (2007), and Darwall (2006) each propose a similar argument.

as expressing demands, for instance. Their expression requires a shared framework of value and practice in order to be intelligible. Pleas to be released from these demands require this shared framework too. So, a moral community is constituted by a set of agents capable of understanding, expressing, and responding to common moral expectations by way of interpersonal emotional reactions. Incapacitation from moral responsibility is therefore the incapacity to participate in a moral community (cf. McKenna 2008a/1998).

With the preceding sketch of Strawson's theory of moral responsibility in place, I turn now to two cases, one really bad and one really good. Using them, I will illustrate the two challenges to the Strawsonian view noted above and also call attention to two reactive emotions on the basis of which I will attempt to defend the view.

## 2. The Really Bad Case

It is natural to think that acts of evildoing are appropriately addressed by the reactive attitudes expressive of blame. Consider the following case:

**Harris:** Robert Harris and his brother planned to rob a bank. They hijacked the vehicle of two 16-year-old friends. Robert promised to leave the friends some money in the car for having used it in the robbery. The car ride was amiable. One of the teenagers wished Robert "good luck." As the teenagers left, Robert raised the rifle and shot one of them in the back. He chased the second down a hill, shooting him four times. He knelt over him and shot him in the head. Harris proceeded to laugh. Harris and his brother left the scene and he began to eat the teenager's lunch 15 minutes after the murders. Harris offered his brother some food. He refused—nauseated from the murders. Harris mocked his weakness. He proceeded to suggest that they impersonate police officers and inform the teenagers' families of the deaths. He went on to suggest that they drive near the scene of the murder in order to kill police officers. Harris, during final preparations for the bank robbery, noticed blood on the rifle and said, "I really blew that guy's brains out." He again began to laugh.<sup>6</sup>

---

<sup>6</sup> This is paraphrased from Miles Corwin's *Los Angeles Times* article, "Icy Killer's Life Steeped in Violence," as quoted in Watson (2008/1987)

How do people respond to Harris? By all appearances, he is clearly blameworthy if anyone is! In prison, he continued to be a man who cared for nothing and no one. His fellow inmates on death row planned to have a party upon his death (Watson 2008/1987, 125).

Yet, as Watson points out, Strawsonianism seems to suggest that Harris is *ineligible* for moral responsibility. How so? Harris expresses unfathomably ill will by showing no consideration for others. He communicates with us. Yet, as Watson points out, “not all communication is dialogue;” Harris repudiates the moral community and places himself outside of it (2008/1987 128). His being outside the moral community is precisely what makes him so evil. His evil consists in part of the fact that our reactive attitudes towards him are pointless—at best they are met with “icy silence” and at worst with “murderous contempt” (ibid. 128). Harris is not a suitable target for our communicating our reactive attitudes—or so it seems. Due to the structure of the Strawsonian theory, Harris appears to be exempt from morally responsibility. But if he is exempt, it is not because of a lack of interpersonal capacities. If he is exempt, it is because extreme evil is its own kind of exemption, and this is absurd.

I have just gone through the first problem: extreme evil ends up being an exempting condition from moral responsibility. There is another problem. Consider Harris’s childhood:

**Harris, cont’d:** Robert Harris was born prematurely. His father had kicked his mother in the stomach after accusing her of infidelity. Robert spent the first months of his life in an incubator at the hospital. Robert’s father beat and abused all his children and his mother. He sexually assaulted his daughters. He never accepted Robert as a son. Over time, his mother grew to blame Robert for her own abuse and came to hate him. Later, she would say that she felt that Robert’s crimes were her fault. She was never able to love him. His sister reported that Robert was starved for attention. He would seek out his mother’s touch, only to be kicked away. She went on to report that he was the most sensitive of all 10 siblings. As a child, he cried when Bambi’s mother was shot when watching the eponymous film. Robert suffered from learning disabilities and was teased at school. At age 14, Robert was sent to a youth detention

center for stealing a car. He was raped several times and attempted suicide twice.<sup>7</sup>

The process from heartbroken child to heartless murderer was brutal. Now it no longer seems clear that it is appropriate to blame Harris, at least not without significant attenuation. Watson describes our reaction to Harris as a mixture of sympathy for the child and antipathy towards the man, leading to ambivalence. This clash is accompanied by the unsettling thought that “one’s moral self is such a fragile thing. One tends to think of one’s moral sensibilities as going deeper than that (although it is not clear what this means). This thought induces not only an ontological shudder, but a sense of equality with the other: I too am a potential evildoer.” (2008/1987, 132) Here is the second problem: Harris was not responsible for his history. No one is—at least not early on. Watson contends, rightly, that we are often ignorant of the historical considerations that shape us. He also discusses possible ways to understand the sensitivity of our reactive attitudes to historical considerations within the Strawsonian paradigm.<sup>8</sup> However, the fact remains that when one focuses his or her attention on histories, one comes to feel like “it is not *one’s* business to blame” (ibid. 137). Faced with Harris’s history, we might think to ourselves, “Who am I to judge? It was only luck that kept me from being as evil as Harris. I was spared such a cruel childhood. If I had been in his position, I might have been no different.” And what is true for you is true for anyone. No one is in a position to blame if we are subject to such deep moral luck. If a history can make us feel like no one is in a position to blame, how can we maintain a commitment that anyone is blameworthy? For a Strawsonian it is impossible to pull apart the propriety of reactive attitudes from

---

<sup>7</sup> Again, this is paraphrased from Miles Corwin’s *Los Angeles Times* article, “Icy Killer’s Life Steeped in Violence,” as quoted in Watson (2008/1987)

<sup>8</sup> What is at issue here is whether or not responsibility requires an “ultimacy” condition, that one be the sole originator of oneself. For the details of Watson (2008/1987)’s suggestions see pg. 133-134 and pg. 137. For a further compatibilist discussion of the ultimacy conditions see McKenna (2008b)

responsibility because the theory says that *to be responsible* is to be the appropriate target of those attitudes. If no one can appropriately blame by addressing other persons with reactive attitudes, then it looks like no one can be blameworthy. The theory appears to be self-defeating.

### 3. A Possible Solution

Consider the following defense of the Strawsonian theory offered by McKenna (2008a/1998). First, we can get the right result with a minimal revision of the basic theory. So long as someone has the capacity to participate in the moral community, then he or she is the appropriate target of reactive attitudes. *Pace* Watson, it is clear that Harris does understand the values of our community. He must in order to mock and repudiate them. It follows that he has the capacity to participate in the moral community in a way that, for instance, a frightening animal does not. So we should hold him blameworthy. Second, the skeptical force of Harris's "ontological shudder" is unwarranted. If the skeptical worry arises from the suspension of our reactive attitudes, we need only admit that the appropriateness of those reactive emotions is distinct from what emotions we in fact have. In the Harris case, we can understand that certain negative emotions are appropriate, even when the emotions we actually experience are sympathetic. If the skeptical worry arises out of concerns for moral luck or equality, then you must realize that given who you are now you could not become like Harris. You lack the potential for evil (unless, of course you are like Harris, but that is a different sort of case). Even if we ought to be compassionate in holding Harris accountable, "the moral order" must find its voice through blame (2008/1998, 217). We owe that to his victims.

Could this be the solution the Strawsonian is looking for? I will argue that it is not. Nevertheless, McKenna's response moves us in the direction of a better reply.

Let's assess McKenna's response to the *reductio*: we can hold those outside of the moral community responsible so long as they have a (here and now) capacity to be members. This response strikes me as insufficient. We must conclude that Harris does indeed have the capacity (here and now) to meet the expectations and demands of morality. Yet, contrary to McKenna, I think that it is in some sense appropriate for us to *stop* engaging Harris with moral demands while maintaining a commitment to his blameworthiness.<sup>9</sup> Why? At least at for some time, he did not care for our indignation. He was unresponsive to moral demands. A perfectly appropriate response to such a person is to *stop trying to get them to respond*. Consider the following case:

**Cheater:** Charles consistently cheats when playing little league baseball. The other little leaguers call him out on his cheating, but Charles does not care. He continues to cheat in spite of this. Eventually, the other children stop trying to penalize Charles. Instead, they just don't let Charles play with them anymore.

By excluding Charles, the little leaguers protect the integrity of the game in a way that is distinct from either acting so as to hold him accountable or exempting him from blame. Our response to the Harris case, and our protection of "the moral order," is more like **Cheater** than we might initially imagine. To block the *reductio* worry, a reactive attitude that expresses this response is needed: Harris is an appropriate target of moral demands, but it is also appropriate for us to withdraw our efforts to communicate those demands to him.

Regarding the self-defeating problem, I agree with McKenna's distinction between the eliciting of a reactive attitude and the understanding of the appropriateness conditions of having a reactive attitude. It successfully replies to the worry about the presence of sympathetic emotions when faced with Harris's history. However, we do ourselves a disservice by

---

<sup>9</sup> On the conversational model I have adopted from McKenna (2012), there is a sense of engagement relevant to the features of one's interlocutor in an instance of moral address. How we decide to engage another person with moral demands ought to depend on what they are like. I will have more to say on this point in the concluding section.

acknowledging that we cannot become like Harris. Still, we would also do ourselves a disservice if, like Watson, we concluded that we could be like Harris. Both responses miss the heart of the “shudder.” Why? Here, my answer is crucial to the view I will advance: *We feel degraded when faced with evil.* This degradation is not directed at any particular feature of oneself—an assumption shared in the disagreement between Watson and McKenna.<sup>10</sup> Rather, it is directed at our common humanity. For instance, there was no special feature of the soldiers of the Einsatzgruppen death-squads that we might use to explain their participation in genocide. One might wonder, if they were able to become the instruments of genocide, might I? Might anyone? And so, one comes to *understand*, not simply wonder, that *we* are able to become evildoers. This challenges our status as moral beings. It is degrading.

Both Watson’s and McKenna’s explanations of our reactions to Harris are tempting because they are explanatorily powerful in cases dissimilar to **Cheater**. Consider this example:

**Insult:** Tom and Tanya are coworkers. Tom, for no good reason, calls Tanya a jerk. Tanya, however, knows that Tom has had a particularly bad day at work—he just lost the big account and it was not his fault.

How might Tanya respond to Tom? There are at least two plausible answers:

**Insult, cont’d 1:** Tanya feels antipathy toward Tom for calling her a jerk for no reason. At the same time, she feels sympathy for him, because Tom’s unfairly losing the big account was a tough break. She remains conflicted.

**Insult, cont’d 2:** Tanya thinks it would be appropriate to resent Tom for calling her a jerk, but she just can’t bring herself to do it. Tom’s unfairly losing the big account through was a tough break, and her current emotional state is just sympathetic towards Tom.

I see no *prima facie* reason to favor either of Tanya’s possible reactions. Local variation in practices and personal temperament license either response. But **Harris** is not comparable to

**Insult.** Tom, we can safely stipulate, would care about Tanya’s reactive attitudes. Harris is, like

---

<sup>10</sup> One could feel degraded for all sorts of particular features of oneself, although here the difference between shame, humiliation, and degradation becomes obscure.

the child in **Cheater**, a target of demands but unmoved by them. What the Strawsonian needs is a different solution to a different kind of case.

#### 4. The Really Good Case

McKenna's account draws our attention to the form of a satisfactory solution to Watson's *reductio* and self-defeating problems. Finding this solution will be easier to reach if we reflect on a different kind of example, the opposite of **Cheater**:

**Paragon:** Patricia consistently follows the rules when playing little league baseball. She also encourages other players to do so. She is furthermore a paradigm example of good sportsmanship. Regardless of how well she plays, the other little leaguers want her on their team. When she plays well, Patricia makes sure to credit the team.

There is an attitude that is appropriate for the little leaguers to have towards Patricia that is distinct from both praise and admiration—one of inclusion—that is the opposite response of the children in **Cheater**—one of exclusion. Let me be clear: Patricia is clearly praiseworthy, but she is the appropriate target of a further attitude (more on this later). Moving to serious matters, consider the following case:

**Williams:** On January 13<sup>th</sup>, 1982, Air Florida flight 90 “crashed into the barrier wall of the northbound span of the 14th Street Bridge” between the District of Columbia and Virginia, “and plunged into the ice-covered Potomac River...Four passengers and one crewmember survived the crash.” they were all located in the same section of the downed airliner, a section of the aft cabin that separated during the crash. A fifth passenger was located in this section of the plane, and was able to get out of the plane and into the water with the other survivors. Although they were not far from shore, ice on the river made rescue by boat impossible, and eventually, a rescue helicopter was sent. The rescue rope from the helicopter was dropped to that fifth passenger, Arland Williams. He did something amazing. He passed the rescue line on to another passenger, and did so repeatedly, until he drowned when the plane wreckage upon which he stood shifted and sank in the water.<sup>11</sup>

---

<sup>11</sup> The preceding is directly quoted from the National Transportation Safety Board, “Aircraft Accident Report: Air Florida, Inc., Boeing 737-222, N62AF, Collision with 14th Street Bridge, Near Washington National Airport, Washington, D.C., January 13, 1982,” and paraphrased from McDougall (2007).

I hope Williams's actions inspire you. He had showed the (arguably) greatest love we can have for other persons by giving up his life to save others. Williams is not simply admirable and praiseworthy. He merits a further response, like the child in **Paragon**.

One powerful reaction to Williams came from the essayist Roger Rosenblatt (1982), who eulogized Williams as "the man in the water." Contemplating Williams's last seconds, he writes:

For at some moment in the water he must have realized that he would not live if he continued to hand over the rope and ring to others. He had to know it, no matter how gradual the effect of the cold. In his judgment he had no choice...

... He was there, in the essential, classic circumstance. Man in nature. The man in the water. For its part, nature cared nothing about the five passengers. Our man, on the other hand, cared totally. So the timeless battle commenced in the Potomac...

Since it was he who lost the fight, we ought to come again to the conclusion that people are powerless in the world. In reality, we believe the reverse, and it takes the act of the man in the water to remind us of our true feelings in this matter. It is not to say that everyone would have acted as he did, or as [the other rescuers] Usher, Windsor, and Skutnik. Yet whatever moved these men to challenge death on behalf of their fellows is not peculiar to them. Everyone feels the possibility in himself. That is the abiding wonder of the story.... If the man in the water gave a lifeline to the people gasping for survival, he was likewise giving a lifeline to those who observed him.

The odd thing is that we do not even really believe that the man in the water lost his fight.... The man in the water pitted himself against an implacable, impersonal enemy; he fought it with charity; and he held it to a standoff. He was the best we can do.

Rosenblatt notes that Williams exemplifies the best of what we are, and so he inspires in us the possibility that we too could be good. Williams reminds us of our human dignity.

Williams *looks* like a paradigmatic example of the praiseworthy in the same way that Harris *looks* like a paradigmatic example of the blameworthy. In order to make the cases analogous, we need to examine his life story:

**Williams, cont'd:** Arland Williams grew up in a small town. His nickname "Chub," was "more about personality than pants size, about being a grinning, gosh-golly, aw-shucks kind of guy who wasn't even riled by everyone calling him Chub." He went through ROTC in high school and was educated at a

military college. A classmate reported the expectations for future officers: "That's an unbreakable code. You go last. Your people go first." Never seeing military action, he instead took a post in the United States for his two required years of service. He then became a banker like his father. Williams spent the next two decades of his life checking the numbers of other bankers. At the time of the accident, he was a bank examiner during a banking crisis and was going through a divorce. During this turmoil, he began seeing his high school girlfriend again. She reported an interesting conversation with Arland. She expressed her expectation of "100 hundred percent" commitment in a relationship. He replied with hesitation. "You have to keep a little for yourself. That's what I've learned."<sup>12</sup>

Williams's actions are explainable by his history. Williams had loved ones. They provided a reason to keep hold of the rope. Yet, his military training prepared him for emergencies. How do we react to Williams now? Watson speaks of the "ontological shudder" we feel upon hearing Harris's backstory. Williams instead inspires a steadying. His story does not diminish our wonder. We too can be good.

Given this, it is odd that the cases have similar features. Both Harris and Williams perform actions that are difficult for us to imagine doing. They are both unreactive to some powerful reasons. Harris is unmoved by moral demands made on him. It is by *immunity* to normal self-regarding moral considerations that Williams inspire us. Given this, Williams seems to be, by the structure of the Strawsonian view, somewhat outside the moral community—or maybe it would be better here to say, beyond it or at its upper regions, a place few of us inhabit. Harris's upbringing and Williams's training seem to explain their lack of receptivity to a certain class of reasons, for Harris moral reasons and for Williams prudential ones. Despite the symmetry in the cases we feel degraded by Harris and empowered by Williams. This demands explanation.

In the next section, I will give one by arguing that there are two specific responses to Harris and Williams. I will go on to argue that these responses should be understood in terms

---

<sup>12</sup> Paraphrased and quoted from McDougall (2007).

of distinctive reactive attitudes, attitudes that have not hitherto been recognized as such in the Strawsonian literature on moral responsibility.

### 5. Elevation and Disgust

The literature on free will and moral responsibility lacks a term for the sort of responses found in **Cheater, Paragon, Harris, and Williams**. Luckily, we have a psychological science friendly to the interpersonal paradigm. Research in positive psychology details a family of “other-praising” emotions. Familiar among them are gratitude, admiration, “appreciation, awe, esteem, and respect” (Algoe & Haidt 2009, 107). Additionally, Jonathan Haidt, among others, propose the existence of an emotion called “elevation,” which is “a pleasurable feeling, sometimes involving warm or pleasant feelings in the chest, that trigger[s] desires of doing good deeds,” involves a “spiritual” or self-transcendent feeling, and promotes altruistic behavior (Algoe and Haidt 2009, 106; cf. Haidt 2003a; Haidt & Morris 2009; Schall, Roper, & Fessler 2010). This builds on the “broaden and build” paradigm of the positive emotions, which hypothesizes that positive emotions expand a person’s “thought-action repertoires” by encouraging exploration, play, and cooperation (Fredrickson 2001, *inter alia*). Elevation animates Rosenblatt’s “The Man in The Water” and it is an appropriate response to Williams.

We have an opposing reaction to Harris. Elevation, on Haidt’s view, is contrasted with “social disgust.” He writes:

In all of its components, elevation appears to be the opposite of social disgust. Where social disgust is caused by seeing people blur the lower boundary between humans and non-humans, elevation is caused by seeing people blur the upper boundary between humans and God (i.e., saints, or people who act like saints). Where disgust makes people close off and avoid contact, elevation makes people open up and seek contact. Where disgust creates negative contamination...elevation creates positive contamination (e.g., people want to touch living saints, or in some cultures to collect the hair, clothing, or bones of dead saints). (Haidt 2003b, 852-879)

One appropriate reaction to Harris is a kind of disgust.<sup>13</sup> The literature on disgust is too expansive to canvass here. Rozin, Haidt, and McCauley (2008) provide one way to frame the issue. They consider disgust to be a scalar defensive response ranging from bodily protection to the protection of the social order. Forms of disgust that seek to protect the body include “distaste” as a defense reaction to poison. “Core disgust” is a defensive reaction to possible contamination from disease. “Animal disgust” is a defensive reaction to death and sexuality—things that might remind us of our animal nature (cf. Rozin and Fallon 1987). Social disgust might be a better name for disgust over offenses of social norms rather than moral offenses. For specifying something relevant to moral responsibility we might choose “moral disgust,” reactions aimed at protecting the moral order itself. I claim that Harris morally disgusts us.

Behaviorally, “disgust is manifested as a distancing from some object, event, or situation, and can be characterized as a rejection;” phenomenologically it is typically manifested as revulsion (Rozin, Haidt, and McCauley 2008, 758-759). This comports well with other theories of disgust. Daniel Kelly (2011) develops a story of how the psychological-physiological disgust system responses that protect the physical body from contamination could be co-opted to protect the “soul” from spiritual defilement. Martha Nussbaum (2004) endorses a similar story in her work on disgust and legal theory. She carefully separates disgust from both anger and indignation as an emotion whose “core idea...is that of contamination of the self; the emotion expresses a rejection of the possible contaminant” (99).

Disgust paradigmatically manifests as an attempt to create *distance* between oneself and the harmful contaminant. This may operate by a kind of magical thinking, which may be a heuristic: once one has contacted the contagion one is always in contact with it and that shared

---

<sup>13</sup> My discussion owes much to Strohming (2014)’s thoughtful guidance through contours of the current debates about disgust. My presentation here follows hers closely. A reader looking for more information about this growing topic should begin with her excellent overview.

properties indicate a shared identity (Strohminger 2014, 483). For instance, people dislike eating chocolate shaped like dog feces (Rozin, Millman, & Nemeroff 1986), and people prefer not to touch objects handled by AIDS patients even after extended periods of time (Rozin, Markwith, & Nemeroff 1992). Objects that share properties with something disgusting—even when it is known that no actually harmful properties are shared—can make an object disgusting. Elevation has the opposite effect. We want to be *closer* to the elevating person.

Once someone like Harris is killed, imprisoned, or banished, the moral order regains stability because the moral contaminant is removed. This is structurally similar to the children’s response in **Cheater**. Animal disgust, reminding us of our animal nature, may make us feel degraded. One thinks, “I am just an animal.” Moral disgust likewise can elicit self-directed degradation. Disgust involves magical thinking. We do share properties with Harris, even if we do not share any harmful properties, and so we feel the “ontological shudder.” We *cannot* conceptually distance ourselves from him. Conversely, Williams elevates us, our response to him is structurally similar to the response in **Paragon**, and we think that we too can be. We *want* to include him in our moral community, even in death.

Disgust is often considered morally dangerous (cf. Kelly 2011, Nussbaum 2004). I am, for now, using the term “moral disgust” descriptively, not prescriptively.<sup>14</sup> As a matter of fact, if someone does not respond to our blame, we may become morally disgusted. I will return to the prescriptive status of disgust.

## 6. Peculiar Reactive Attitudes

Elevation and moral disgust—disgust henceforth—have the right features to be reactive attitudes. Strawson demarcates the reactive attitudes by tying them to our “demand [for] some degree of goodwill or regard on the part of those who stand in...relationships to us”

---

<sup>14</sup> As Giubilini (2015) points out, “moral disgust” is ambiguous between the two uses.

(2008/1962, 23). Elevation and disgust respond to quality of will. As McKenna points out, there is an intelligible quality of will expressed by Harris's actions—a terribly evil one. Williams's quality of will is extraordinarily good. I wish to argue that elevation and disgust are suitable response to these cases but are not themselves praise and blame. However, someone might resist this conclusion by suggesting that elevation and disgust are non-standard forms of praise and blame. Let me be clear: disgust is a response to a blameworthy agent. But, it is a response that has, in some sense, left blame behind. Disgust expresses a judgment that someone should be excluded—not exempted—from participation in the moral community.<sup>15</sup> This is distinct from actively holding someone accountable for his or her actions, but it presupposes that the agent in question is blameworthy. Consider **Cheater** again. The players dismiss the cheater from the game, not because he is ineligible, but because he won't follow the rules. This response presupposes that he is a legitimate rule breaker! Thus, I claim that disgust presupposes our commitment to an agent's blameworthiness, but it itself does not constitute a way of holding someone accountable.<sup>16</sup> We are trying to distance ourselves physically and conceptually from the agent who repudiates morality. Thus, a fitting target of disgust is an agent who repudiates the demands of morality while generally being able to comply with them.

Elevation expresses inclusion in the moral community, and this presupposes a commitment to the praiseworthiness of the agent in question. A fitting target of elevation is

---

<sup>15</sup> I follow R. Jay Wallace (1994) in thinking that reactive attitudes have beliefs about quality of will, involving judgment, as their objects.

<sup>16</sup> The person who considers disgust to be a non-standard form of blame is suggesting something close to Scanlon (2008)'s sense of blame: blame is whatever response is appropriate from one person towards another person who through some action has indicated that they hold attitudes that impair the relationship between them (122-123, 128, 138). On this kind of view, disgust is a form of blame because it is an appropriate response to an impaired relationship. Harris holds attitudes that make it appropriate for us to be disgusted and distance ourselves from him. However, my analysis of disgust indicates that something is very wrong with this view. We maintain our commitment to Harris's blameworthiness, but we take no action towards him that constitutes blame. We *give up trying to express blame him*. To call this response "blame" in the fullest sense is absurd.

an agent who exceeds the demands of morality by showing exemplary quality of will. Their quality of will shows better than expected regard for others often due to a lack of consideration for his or herself. This, on the face of it, seems to be the same conditions for praiseworthiness (cf. Pereboom 2014, 127). Recall that I claimed in **Paragon** that the exemplary little leaguer merited a further response beyond praise. I take it to be an open question whether or not elevation “leaves praise behind,” but I maintain that elevation and praise are *distinct* responses to praiseworthy agents. There could be an asymmetry between elevation and disgust in this way, and we would have a natural explanation for this insofar as praiseworthy agents do not ignore or repudiate our moral address. Still—to take Haidt (2003b)’s example of elevation related behavior quoted above—collecting someone’s bones seems to be a very different sort of thing than praise. I will leave this asymmetry question open but maintain a commitment to the distinctness of elevation and praise.

One might be skeptical that disgust is a reactive attitude because one might think that is not a form of moral address. In her defense of contempt, Macalester Bell argues that disgust is not a form of moral address because it does not make any *demands* on an agent. She writes (2013, 187-88):

Disgust presents its object as a contaminant and therefore as incapable of a response that could be given uptake: Any response that the target of disgust could offer would be unwelcome because contagions should be kept at arm's length. We don't want slime talking back to us (this is the stuff of horror movies!), nor do we want those we deem disgusting responding to our evaluation of their disgustingness. Moreover, if the message implicit in disgust is that its object is a contaminant that must be avoided, it is difficult to see how this message could be given uptake because it is not clear what would count as taking this claim seriously.

Bell is right to see disgust as a problematic bearer of moral demands in the same way that resentment, indignation, or contempt are such bearers. But this does not mean that disgust fails to be a form of moral address. When we are morally disgusted, we have *already* addressed

demands to an agent and they have failed to meet them. But they also continue to ignore our demands. By expressing disgust, we communicate a powerful message: we won't deal with you anymore. And this is a kind of demand, in the sense of an ultimatum. Moreover, we can get over our moral disgust of someone when they let us know that they recognized our demands, and this suggests that we really were communicating after all. So, I do not think we should cede to Bell that disgust isn't communicative of a demand and therefore not a kind of moral address.<sup>17</sup>

Even if disgust and elevation are distinct from praise and blame, one might think that their distinctiveness is accounted for by being aretaic or characterological responses to agents.<sup>18</sup> Here is one clear way of spelling this out: David Shoemaker (2015) has recently defended a theory of responsibility whereby different kinds of emotional reactions pair with different aspects of "quality of will." On this view, emotions expressive of accountability, like resentment and indignation, are tied to someone's regard for others. However, emotions like admiration and contempt are tied to someone's character and are expressive of another dimension of responsibility: attributability (cf. Watson 1996). Why not say that we admire Williams and have contempt for Harris? Does this explain the cases?

I am not convinced. The cases involve the maintenance of a commitment to moral responsibility in the accountability sense, and this demands a solution that focuses on accountability. Harris and Williams seem blameworthy and praiseworthy for their *actions* respectively, and yet they are inapt candidate for moral address by way of reactive attitudes. This is why the cases are puzzling. Contempt and admiration are evaluations of character.

---

<sup>17</sup> It shouldn't surprise us that disgust at slime is non-communicative. Slime lacks the relevant capacities to understand the meaning of reactive attitudes. Disgust is a primitive emotion, and like other moral emotions it only gains moral significance as it becomes a part of our moral practices and becomes interconnected with our moral concepts.

<sup>18</sup> Thanks to David Shoemaker for pressing me on this point in conversation.

They are therefore unsuitable explanations of the puzzle found in the cases I have been considering. However, the deeper worry seems to be that disgust and elevation are not what I say they are, namely, reactive attitudes tracking regard for others in actions.

Let's start with elevation and Williams. Williams is in many respects utterly ordinary. His character is not obviously admirable. He explicitly endorsed some selfishness in his relationships (recall his comment about not committing oneself fully). It is precisely because his character is not extraordinary that we find his action so extraordinary. I claim that *he* elevates us as an exemplar of what *we* are capable of. His character may modify our response to him, but he is not his character. His character may provide a backdrop by which our responses to him are modified. This modification may be especially apparent in cases where evildoers have (as Kant would call it) a "revolution of the will" and perform an extremely good action (cf. Kant 1998/1794 6:48). In cases like this, we cannot admire the evildoer's character (it's evil). But it precisely because they are an evildoer that their sudden regard for others shines all the brighter.

On to disgust: I have inherited the Harris case, and admittedly it is harder to pull apart his character and his regard for others in his actions. But disgust and contempt are distinct in the following case:

**Cheater II:** Cassandra and Carol have been friends and coworkers for 20 years. They work in public service. Cassandra has always looked up to Carol and admired her for her hard work and dedication. However, Cassandra finds out that for the past 10 years ago, Carol has been taking bribes. Cassandra confronts Carol, who thinks that accepting the bribes is morally justified. Cassandra thinks it is wrong but cannot do anything more; her evidence for the bribery would not convince anyone else. Cassandra genuinely believes that Carol is a good person who has made a mistake in her moral reasoning. But she cannot help distancing herself from Carol, feeling a sense of revulsion at the thought of her actions.

Cases like this show that we can separate out disgust and contempt.<sup>19</sup> Cassandra does not have contempt for Carol, even for her habitual transgressions. Instead, she feels “yucky” about her longtime friend. In light of a good character, especially transgressive actions—a dedicated public servant taking bribes, for instance— can make disgust feel especially pronounced.<sup>20</sup>

There is something to be said about why the aretaic reading of disgust and elevation is so tempting. Strawson gives the reactive attitudes a tripartite division: attitudes in reaction to another’s quality of will towards oneself, vicarious attitudes in reaction to another’s quality of will towards another person, and self-reactive attitudes related to the demand on oneself to show others good will (2008/1962, 28-29). Disgust and elevation can be other-directed *and* self-directed *and* vicarious. In the above I described their other-directed features. As they are self-directed, they share features with other self-directed emotions like shame and guilt. Arguably, guilt as a self-directed emotion is a response to a specific feature or action, whereas shame is a global negative self-assessment. It is the difference between saying, “I did a terrible thing, I ‘X’d’” and “I ‘X’d’ and therefore I am terrible” (cf. Lewis 2008). Elevation and disgust are like shame. Notice that praise and blame *lack* these self-directed features. For instance, when I am indignant, I direct no attitude towards myself. Elevation and disgust are also vicarious attitudes. Remember, the “magical thinking” of disgust: shared properties imply shared identity. Not only us, but also *everyone*, is judged to be bad in the same way. Elevation seems to involve similar, but positive, thinking.

A *global* evaluation is not necessarily an *aretaic* evaluation, but aretaic appraisals are often global. We can be interested in a specific feature of someone’s character and we can be

---

<sup>19</sup> It also shows that the exclusionary effect of disgust need not be total. More on this later.

<sup>20</sup> Empirically, contempt and disgust appear to have different (though related) characteristic and universally recognizable facial expressions. This suggests that they are distinct basic emotions (Ekman and Friesen 1986, as cited by Bell 2013).

interested in his or her character as a whole. And if the above cases are right, there can be global evaluations that are distinct from evaluations of character. For instance, consider Williams again; his life was not structured around selfless cares and commitments for others. His character is in many ways ordinary. He just performed one all-consuming act, and this makes us *see* him as the man in the water. That epitaph is, if anything, a global evaluation. In a single moment he was *for the good* in some fundamental sense. His character and his history just drop out of our picture of him in the face of his regard for the lives of others. Their disappearance is part of what makes the action so elevating—some actions make us feel *unbound* by who we are and where we come from. Disgust is structurally similar—I'll elaborate in the next section.

I contend that elevation and disgust are reactive attitudes. They are emotional responses to regard or lack thereof for others as expressed in intentional action that are distinct from, if intimately related to, evaluations of character. My suggestion that we accept them as reactive attitudes is an unorthodox amendment to Strawsonian views. I will now argue that their acceptance can solve the two problems. The amendment is warranted.

## 7. Solutions

Strawsonians need an account of why Harris is the appropriate target of reactive attitudes generally *and* why it is appropriate to stop addressing him with moral demands. This is the form of the solution to the *reductio* and self-defeat problems.

So, how can we maintain our commitment that Harris is blameworthy, but why is it also appropriate for us to stop engaging him with moral demands? He is generally an appropriate target of blame because he can understand the values of the moral community. Yet, it is not appropriate to hold him morally accountable because he repudiates our moral demands. I have argued that Harris is the appropriate target of disgust. With disgust comes a

defensive reaction against that agent as a moral contaminant. Our response to Harris tries to re-establish the purity of the moral-interpersonal world by exclusion. In this case, permanent exclusion: Harris was sentenced to death.

The foregoing implies that Harris is not exempted. Like Strawson, I maintain that exemption involves viewing someone from an objective attitude. Exemption means that someone is not the sort of thing that can be addressed by reactive attitudes. In fact, it would be *easier* to view Harris objectively. “It was just his childhood!” we might tell ourselves. But thinking that Harris ought to be exempted misunderstands the *personal* nature of the case and his crimes. Given this, the Strawsonian who accepts disgust avoids the *reductio*. It is not absurd to think that Harris elicits in us a complex response that is not blame, but is also not exemption, and is instead exclusion. What about Williams? *Exceeding* the criterion for responsibility should not exempt someone from responsibility. I agree. He is an apt target of praise, but also the distinct attitude of elevation.

Given disgust, the Strawsonian need not fear the self-defeat problem. We do make a negative self-assessment when faced with Harris that involves some generalization about our common nature. This generalization, however, is not one about our particular histories, and this was the source of the self-defeat problem. I have argued that elevation and disgust do not focus on character or history primarily, even if they are global responses to others. Moreover, even if we feel sympathy towards Harris because of his past, this does not dispel our disgust. Why? History and character can *heighten* our feeling of disgust even in the presence of sympathy. Like how violations of norms about sex and death can remind us of our animal nature, Harris’s history reminds us of our own moral fragility. Disgust makes us feel *bound* to who we are and where we came from. Insofar as Harris makes us feel this way, we are morally humiliated.

But for every case that instills doubt about ourselves, we can find another to restore our faith, like the case of “The Man in the Water.” The lesson to learn from this is that the moral community has no well-defined boundaries. Persons sometimes fall above and below the threshold of ordinary moral expectation. Sometimes this can happen by an agent’s own doing. Like Harris, persons can remove themselves from the moral community. Internal to our practices there are reactive attitudes that express our responses to such persons. When we see someone who rejects moral expectations, we exclude them. When we see someone do *better* than those expectations, we raise the expectations we have on ourselves.

Strawson himself has more to say. He considers “something far above or far below the level of common humanity,” i.e., something that breaks free of our “common roots in our human nature and our membership in human communities.” He calls the thing that lacks the self-directed attitudes a saint and the thing that lacks the other-directed attitudes a moral idiot. We imagine Harris to be a moral monster. But Williams and Harris are not Strawson’s imagined beings even if they seem like them. As Watson reports: “in his last years, Harris either remained, or became once again, capable of friendship and remorse. His crimes were monstrous, but he was not a monster after all. He was one of us” (Watson 2008/1987, 141). Why does Watson tell us this? Harris mouthed the words “I’m sorry” to one victim’s father moments before he was put to death by the state (ibid, 140). So, he was not *really* a monster. Given this, maybe disgust is immoral. Should someone like Harris really face fatal exclusion from the moral community?

## 8. The Ethics of Disgust

The answer to the preceding question is difficult. Here is the problem. Disgust is my way out of a theoretical problem for the Strawsonian responsibility theorist. But what if disgust itself is morally problematic? Many authors have worried about the morality of disgust as a

response to human persons. It often misguides our moral lives. Immigrants, people of different sexual orientation, race, and so on, have both presently and historically been the targets of unmerited disgust. For instance, Daniel Kelly is skeptical of the possibility of genuine moral disgust because disgust has become associated with immoral attitudes (2011, 128). And Martha Nussbaum is positively anti-disgust. Indeed, one might say that she is disgusted by disgust. She writes: “I would argue...that even the moralized form of disgust is an emotion that is highly problematic. It must be contained and perhaps even surmounted, on the way to a genuine and constructive social sympathy” (2004, 105).

The disgust skeptic has a serious worry. Consider this argumentative strategy: Bell (2013) argues (to my mind, persuasively) that contempt can be both a fitting and reasonable response to persons who embody the “vices of superiority”, like arrogance or hypocrisy. Moreover, holding those persons in contempt is the best moral response to those vices available. We can say similar things regarding disgust. Moral disgust can be fitting towards persons who are genuinely unmoved by moral demands. It can be reasonable to be disgusted if we have sufficient evidence that the target of our disgust is truly unmoved. Although I do not endorse moral disgust as the required response to the moral fault of being unmoved by calls for accountability to the demands of morality, I do think disgust is uniquely good as a response to this fault.

This will not convince the disgust skeptic. She may object to this view in several ways. For instance, disgust skepticism can be motivated by disgust’s exclusionary effects. On the view I have offered, disgust has the functional role of protecting the moral community by excluding persons who are not open to moral address. This may seem unduly harsh, especially in light of Watson’s post-script to “Responsibility and the Limits of Evil”. This problem can be resolved if we endorse the view that the exclusion brought about by disgust comes in

degrees. And we should endorse that view. Consider again **Cheater II**, in which Cassandra is disgusted by Carol's corruption in the public service. Cassandra does not think that Carol should be excluded from the moral community, full stop. (She *should* not, in any event). Cassandra nevertheless seems warranted in distancing herself from Carol. What Carol has done *is* morally disgusting. It is not as morally disgusting as eating the lunches of one's murder victims. In extreme cases of evildoing, the appropriate degree of exclusion may indeed be removal from the moral community. The degree of exclusion which a morally disgusting act merits therefore seems commensurate to the degree to which the action itself merits disgust.<sup>21</sup>

Let's consider a different way to motivate disgust skepticism. One might wonder how disgust can ever be a fitting response to a human person if one thinks that disgust is inherently dehumanizing. Consider that dehumanization is a possible outcome of the Strawsonian view of exemption anyways. From an objective attitude, another person is as an object of casual or explanatory inquiry. We ask, "how it is structured and/or how it functions" (Bennett 2008, 53). The objective attitudes at their most extreme—although often accompanied by positive non-reactive emotions—can dehumanize persons by reducing them to non-agents (cf. Strawson 2008/1962, 24). That the objective attitude is dehumanizing is consistent with Strawson's description of it as reducing human behavior to causal understandings of "programmed mechanisms" (Strawson 2011, 140). These kinds of explanations, at least some of the time, can be morally degrading. Despite this, we do not think that exemption is never a fitting response to human persons. Likewise, we are disposed to experience disgust. It is surely true that some token instances of disgust in some domains, like those Kelly and Nussbaum

---

<sup>21</sup> I offer no calculus here; in real life cases, the degree of appropriate disgust, and therefore appropriate exclusion, will likely be a matter of good moral judgement.

find objectionable, are highly problematic and unfitting. But it does not follow that all tokens of disgust are.

I think we should be unmoved by the worry that disgust is, in itself, morally problematic because it is dehumanizing. But even if disgust can be an appropriate moral response, perhaps disgust is generally problematic in some other way. Disgust is an admittedly blunt instrument in morality's toolbox; one might worry that it promotes bias in moral judgement.<sup>22</sup> If we are frequently disgusted, we may be more likely to be disgusted by all the wrong things.

The worry appears empirically supported. In a series of interesting experiments, Jones and Fitness (2008) found that individuals who were highly sensitive to disgust were prone to several biases in their judgment.<sup>23</sup> Jones and Fitness believe that this is evidence for the view that individuals who are highly sensitive to disgust exhibit what they call *moral hypervigilance*, “a range of cognitive, emotional, and behavioral tendencies that are intended to reduce the risk of exposure to moral contaminants” (2008, 614). They found that, for instance, persons who were highly sensitive to disgust made higher estimates of the probability that suspects described in crime vignettes were culpable and were more likely to recommend lengthy sentences for criminals. Moreover, persons who are sensitive to disgust also had comparatively inflated perceptions of community crime levels. Perhaps most interestingly, Jones and Fitness found that highly disgust-sensitive persons also displayed a greater tendency to attribute evilness to criminals. These findings fit well with my view that moral disgust presupposes accountability but involves a distinct exclusionary response especially aimed at evildoing. This

---

<sup>22</sup> Thanks to an anonymous reviewer for raising this important objection and pointing me towards the relevant empirical research.

<sup>23</sup> Jones and Fitness used versions of the disgust scale presented in Haidt et al. (1994) to measure disgust sensitivity. Importantly, this disgust scale does not specifically focus on one kind of disgust, e.g., moral disgust.

seems all the worse for my view, though. Persons who were sensitive to disgust displayed systematic differences in their moral judgment than persons who were less sensitive to it. And these judgements were harsh. Disgust seems to lead to hypervigilance against moral contamination, and correspondingly, a tendency towards drastic and potentially unwarranted countermeasures.

In response to this worry, it should first be noted that Jones and Fitness themselves state that their use of the term “bias” is not meant to imply that persons who are highly disgust-sensitive are making errors in moral judgment (2008, 615). They did not measure the accuracy of the judgements in their experiments. They instead use the term to refer to dispositions to make judgements of a certain sort. But let’s assume that highly disgust-sensitive individuals are making moral mistakes. What then? Jones and Fitness do not investigate the relationship between the emotional *state* of disgust and moral hypervigilance. Rather, they investigate the relationship between moral hypervigilance and a personality *trait*, disgust sensitivity.<sup>24</sup> So, even if high sensitivity to disgust does produce problematic bias in moral judgment and hypervigilance, this does not mean that any token state of being disgusted is morally problematic. On my view, moral disgust has an important moral role to play in policing the boundary of the moral community. It is, however, a weapon of last resort. It should be used sparingly and only after calls for accountability via blame fail. Given this, it is perfectly possible that persons who are highly sensitive to disgust may misuse this powerful moral tool and so become inappropriately hypervigilant. Strawsonians take the moral emotions to play a key role in our moral practices of accountability. We do not endorse the stronger thesis that we ought

---

<sup>24</sup> Specifically, Jones and Fitness label disgust sensitivity a trait-like construct that describes the ease and intensity with which one is prone to experience disgust (2008, 614).

to inculcate in ourselves dispositions for hard feelings. We may even say that doing so is morally wrongheaded if these dispositions promote moral error.

To see why such a response should convince, compare with anger. Strawsonians take forms of anger (resentment and indignation) to express (or constitute) blame. That is, *states* of anger express blame. But *trait* anger, being disposed to respond to situations with anger, seems correlated with biases in moral judgement. Epps and Kendall (1995) showed that high trait anger in adults is associated with hostile interpretation bias—the tendency to interpret others as having hostile intent when social cues are ambiguous. Wenzel and Lystad (2005) found that although both anxious and angry individuals demonstrate negative interpretation biases when asked about ambiguous but potentially threatening scenarios, the bias is more extensive among angry people. In Hazebroek et. al (2001)’s study, individuals with high trait anger blamed a wrongdoer more than individuals with low trait anger when presented with videos of social interactions. This difference was especially striking in cases where the wrongdoer’s intent was ambiguous.

Do these results mean that Strawsonians should accept that anger is morally problematic? No. Nothing so far shows the accuracy of the judgements in question. And even if these biases are morally problematic, it only shows that angry people are prone to moral error. Reflection on ordinary life makes this apparent enough. Sometimes, though, we need anger to stand up for what is right and to hold wrongdoers to account. We should say similar things about disgust. We may need it to respond to those who are unresponsive to morality’s demands, in spite of the danger. We should try not to be characteristically angry or disgusted people.

Briefly put, we have need for the more difficult moral emotions even if they come with the risk of error in moral judgment. The lesson to draw about disgust and anger from this is simple: *use with caution*.

## 9. Conclusion

I have argued that two challenges to Strawsonian compatibilism fail. The first was a *reductio* problem that evil agents are exempted from moral responsibility. The second was a self-defeat problem that a person's history dispels reactive attitudes. My response was to suggest the adoption of elevation and disgust as reactive attitudes. This solution, although with questions unanswered, is plausible and explanatorily powerful. It is not an ethically problematic solution either. But beyond contributing a novel solution to these problems, I hope to have shown the theoretical fruitfulness of being true to our experience as moral agents. Our moral reactions to others, especially in extreme cases, are varied, subtle, and messy. This is a feature of our lives to be welcomed rather than ignored.

### Works Cited:

- Algoe, S.B. & Haidt, J. 2009, "Witnessing excellence in action: the "other-praising" emotions of elevation, gratitude, and admiration", *The Journal of Positive Psychology*, vol. 4, no. 2, pp. 105–127.
- Bell, M. 2013, *Hard Feelings: The Moral Psychology of Contempt*. Oxford University Press, New York.
- Bennett, J. 2008, "Accountability (II)" in *Free Will and Reactive Attitudes: Perspectives on P.F. Strawson's "Freedom and Resentment"*, eds. M. McKenna & P. Russell, Ashgate, Burlington, VT; Farnham, England, pp. 47.
- Bennett, J. 1980. "Accountability." In Zak van Straaten, ed. 1980. *Philosophical Subjects: Essays Presented to P.F. Strawson*. Oxford: Clarendon, pp. 14-47.
- Darwall, S.L. 2006, *The Second-Person Standpoint: Morality, Respect, and Accountability*, Harvard University Press, Cambridge, Mass.
- Elkman, P. and W. Friesen, 1986, "A New Pan-Cultural Facial Expression of Emotion." *Motivation and Emotion*, vol. 10, no. 2, pp. 159–168.
- Epps, J. and P.C. Kendall, 1995, "Hostile Attributional Bias in Adults." *Cognitive Therapy and Research* 19, no. 2, pp. 159-178.
- Fischer, J.M., and Ravizza, M. 1998. *Responsibility and Control: An Essay on Moral Responsibility*. Cambridge: Cambridge University Press.

- Fredrickson, B.L. 2001, "The Role of Positive Emotions in Positive Psychology: The Broaden-and-Build Theory of Positive Emotions", *American Psychologist*, vol. 56, no. 3, pp. 218-226.
- Giubilini, A. 2015, "What in the World is Moral Disgust?", *Australasian Journal of Philosophy*, vol. 95, no. 2, pp. 227-242
- Haidt, J. 2003a, "Elevation and the positive psychology of morality." in *Flourishing: Positive psychology and the life well-lived.*, eds. K. CLM & J. Haidt, American Psychological Association, , pp. 275–289.
- Haidt, J. 2003b, "The moral emotions" in *Handbook of affective sciences*, eds. R.J. Davidson, K.R. Scherer & H.H. Goldsmith, Oxford University Press, Oxford, pp. 852-870.
- Haidt, J. & Morris, J.P. 2009, "Finding the Self in Self-Transcendent Emotions", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 19, pp. 7687-7688.
- Hazebroek, J.F., K. Howells, and A Day, 2001, "Cognitive Appraisals Associated with High Trait Anger." *Personality and Individual Differences*, vol. 30, no. 1, pp. 31-45.
- Kant, I. 1998, "Religion Within The Boundaries of Mere Reason" in *Religion Within The Boundaries of Mere Reason: And Other Writings*. Ed. Allen Wood and George di Giovanni. Cambridge: Cambridge University Press.
- Kelly, D. 2012, *Yuck: The Nature and Moral Significance of Disgust*, MIT Press, Cambridge, Mass.
- Lewis, M. 2008, "Self-Conscious Emotions: Embarrassment, Pride, Shame, and Guilt" in *Handbook of emotions*, eds. M. Lewis, J.M. Haviland-Jones & L.F. Barrett, 3rd Edition edn, Guilford Press, New York, NY.
- Macnamara, C. 2015, "Blame, Communication and Morally Responsible Agency" in *The Nature of Moral Responsibility*, eds. R. Clarke, M. McKenna & A. Smith, Oxford University Press, New York.
- McDougall, C. 2007, *The Hidden Cost of Heroism*. NBC News. Rodale Inc, Emmaus.  
<http://www.nbcnews.com/id/21902983/ns/health-behavior/t/hidden-cost-heroism/#.V9Bp07Ut2Uc>
- McKenna, M. 2012, *Conversation and Responsibility*, Oxford University Press, Oxford.
- McKenna, M. 2008a, "The Limits of Evil and the Role of Moral Address: A Defense of Strawsonian Compatibilism" in *Free will and reactive attitudes: perspectives on P.F. Strawson's "Freedom and resentment"*, eds. M. McKenna & P. Russell, Ashgate, Burlington, VT; Farnham, England, pp. 201.
- McKenna, M. 2008b, "Ultimacy & Sweet Jane." In Nick Trakakis and Daniel Cohen, eds., 2008. *Essays on Free Will and Moral Responsibility*. Cambridge Scholars Publishing, pp. 186-208.
- Nussbaum, M. 2004; 2009, *Hiding from Humanity: Disgust, Shame, and the Law*, Princeton University Press, Princeton, N.J.
- Pereboom, D. 2014, *Free Will, Agency, and Meaning in Life*. New York: Oxford University Press.
- Pereboom, D. 2001, *Living Without Free Will*. Cambridge: Cambridge University Press.
- Rosenblatt, R. 1982, *The Man In The Water*, Time Incorporated, New York.
- Rozin, P., Haidt, J. & McCauley, C.R. 2008, "Disgust" in *Handbook of Emotions*, eds. M. Lewis, J.M. Haviland-Jones & L.F. Barrett, 3rd Edition edn, Guilford Press, New York, NY, pp. 757-776.

- Rozin, P., Markwith, M. & Nemeroff, C. 1992, "Magical Contagion Beliefs and Fear of Aids", *Journal of Applied Social Psychology*, vol. 22, no. 14, pp. 1081-1092.
- Rozin, P. & Fallon, A.E. 1987, "A Perspective on Disgust", *Psychological review*, vol. 94, no. 1, pp. 23-41.
- Rozin, P., Millman, L. & Nemeroff, C. 1986, "Operation of the Laws of Sympathetic Magic in Disgust and Other Domains", *Journal of personality and social psychology*, vol. 50, no. 4, pp. 703-712.
- Russell, P. 2004, "Responsibility and the Condition of Moral Sense." *Philosophical Topics* 32: 287-306.
- Russell, P. 1992, "Strawson's Way of Naturalizing Responsibility." *Ethics* 102: 287-302.
- Scanlon, T. 2008, *Moral Dimensions: Permissibility, Meaning, Blame*, Belknap Press of Harvard University Press, Cambridge, Mass.
- Shoemaker, D. 2015, *Responsibility from the Margins*. Oxford, UK: Oxford University Press.
- Shoemaker, D. 2007, "Moral Address, Moral Responsibility, and the Boundaries of the Moral Community", *Ethics*, vol. 118, no. 1, pp. 70-108.
- Smilansky, S. 2000, *Free Will and Illusion*, Clarendon Press: Oxford.
- Strawson, G. 1986, *Freedom and Belief*, Oxford: Clarendon Press.
- Strawson, P.F. 1962/2008, "Freedom and Resentment" in *Free will and reactive attitudes: perspectives on P.F. Strawson's "Freedom and resentment"*, eds. M. McKenna & P. Russell, Ashgate, Burlington, VT; Farnham, England, pp. 19.
- Strawson, P.F., Strawson, G. & Montague, M. 2011, *Philosophical writings*, Oxford University Press, Oxford.
- Strohming, N. 2014, "Disgust Talked About", *Philosophy Compass*, vol. 9, no. 7, pp. 478-493.
- Wallace, R.J. 1994, *Responsibility and The Moral Sentiments*, Harvard University Press, Cambridge, Mass.
- Watson, G. 1987/2008, "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme" in *Free will and reactive attitudes: perspectives on P.F. Strawson's "Freedom and resentment"*, eds. M. McKenna & P. Russell, Ashgate, Burlington, VT; Farnham, England, pp. 115.
- Wenzel, A. and C. Lystad, 2005, "Interpretation Biases in Angry and Anxious Individuals." *Behaviour Research and Therapy* vol. 43, no. 8, pp. 1045-1054.
- Wolf, S. 1981, "The Importance of Free Will", *Mind* vol. 90, pp. 386-405.