Empathic Control?

David Shoemaker[*]
dws267@cornell.edu

ABSTRACT

It has long been thought that control is necessary for moral responsibility. Call this the *control condition*. Given its pride of place in the free will debate, "control" has almost always been taken to be shorthand for *voluntary* control, an exercise of choice or will. Over the last few decades, however, many have been arguing for including a range of attitudes for which we seem to be responsible that, if controlled at all, must be controlled via a very different mechanism, namely, evaluative judgment. Call this second type of control *evaluative* control. In this paper I will present and discuss in detail an additional agential stance – reasonish regard – for which we treat one another as responsible, but that is ungoverned by either of the first two types of control. If we want to require a control condition for responsibility, then, we will need to introduce and include a third type of control, what I call *empathic control*.

> [A] unifying requirement on moral responsibility is
> that control comes in *somewhere*.
> Michael McKenna (2008)

1. Introduction

It has long been thought that there is an inexorable link between agential control and moral responsibility. Indeed, as McKenna stresses in the epigraph, moral responsibility is typically viewed as unified by a *control condition*, according to which a responsible agent "must *control* his behavior in a suitable sense, in order to be morally responsible for it" (Fischer and Ravizza 1998: 13; emphasis in the original). This is also often taken to be precisely what free will consists in, namely, "the unique ability of persons to exercise control over their actions in the fullest manner necessary for moral responsibility" (McKenna, 2012: 188; see also Mele, 2006: 17; Pereboom, 2014; Levy, 2014: 109; and McKenna and Pereboom, 2016: 6).

[*] Cornell University, USA.

Most theorists have taken "control" to refer exclusively to the exercise of volition, a function of the voluntary. Call this view *voluntarist*.[1] While there is some dispute about what the voluntary is, what voluntarists do tend to agree on is that the domain of voluntary control is exclusively over *intentional actions*, where, importantly, these may include mental actions (McKenna, 2008: 30-31). Recent work poses a serious challenge to this approach, though: Agents seem quite morally responsible for *attitudes* over which they nevertheless exercise no voluntary control. If we want to maintain the unifying control condition for responsible agency, then, it seems we need to introduce and incorporate a different type of control to account for our moral responsibility for these non-voluntary attitudes. Control can still be necessary for moral responsibility, but it must come in two flavors, we might say, depending on its agential target (actions or attitudes).

In the first part of the paper, I will set up this dialectic in more detail, and then I will show that allowing this second type of control into the mix introduces a wedge for including a *third* type of control, given that we also seem to be responsible, at least sometimes, for certain ways of *perceiving* others, where this is ungoverned by the previous two types of control. If we must have control over these perceptual stances for them to count as things for which we are responsible, then they will have to be governed by what I will call *empathic* control. After providing some details of what such control would look like and then showing why it really can't be accounted for with the control-based tools already in hand, I will explore the advantages and disadvantages of including it into our control pantheon. This will ultimately be a paper, then, about some tough methodological choices we theorists have to make.

2. From Voluntary to Evaluative Control

Why think we should include a control condition on responsibility in the first place? (I hereby drop the "moral" when talking about responsibility, not only for ease of reference, but also because any plausible theory of responsibility has to apply to all normative domains.) The leading argument is that if blaming an agent for doing something is to be appropriate, just, fair, or deserved, then the blamed agent must have had an opportunity to avoid that blame, which is most often taken to mean that the agent had to have had some sort of control over the doing

[1] Following McKenna 2012 (188).

of that thing (see, e.g., Glover 1970: 73; Nagel 1979: 25; Wallace 1994: 86-87; Kane 1996: 83; Fischer and Ravizza 1998; Levy, 2005; Nelkin, 2011: Ch. 2; McKenna, 2012: 187-88; Pereboom, 2014).[2] This is because blame is typically taken to involve sanction or harm of some sort, and this surely requires moral justification (McKenna, 2012: Chs. 6-7; Pereboom, 2014: Chs. 5-7). If one had no control over doing something, though, it is widely held that being sanctioned or harmed for doing it would lack moral justification.[3]

What are the arguments for *voluntarism* about control? As McKenna (2012: 189) notes, there aren't many. It likely has its roots in Aristotle's famous discussion of the voluntary in the *Nicomachean Ethics*. But there are plausible arguments available for it. One is that voluntarism explains our practices of excuse and exemption quite well (McKenna, 2012: 189). We tend to excuse agents from blame for bad actions that are accidental, coerced, or performed under duress. What unites these excuses, one might say, is that they are all cases in which voluntary control is absent.[4] Similarly, one might view a crucial *exempting* feature – what excludes agents from the responsibility domain altogether – as being an *incapacity* for voluntary control.

The nature of the voluntary is somewhat obscure, although most agree it has to do with deliberate choice and its execution (e.g., Arneson, 2004; Levy, 2014; Searle, 2001: 50, e.g.). All that matters for our purposes, however, is that all voluntarists agree on the object of voluntary governance, namely, *actions*, either physical or mental. What I (deliberately) choose and have direct voluntary control over are things that I *do*. And as what I may justifiably be punished for are only my actions, it would seem, there is a very natural connection between voluntarism and responsibility.

[2] Responsibility for omissions is a tough nut to crack, but there are ways to do it while maintaining a control/freedom requirement. As Randolph Clarke has recently put it, "an agent can freely not do a certain thing" (Clarke, 2014: 104), where by "freely" he has in mind the exercise of some sort of voluntary control.

[3] Here is a point at which Gary Watson's remarks in "Two Faces of Responsibility" are helpful. In discussing responsibility's accountability face, he notes that the fairness of holding people to account (through, e.g., sanctions) requires avoidability, but that avoidability doesn't necessarily entail control over the thing for which one is accountable. You might be hired to maintain order at a dance hall but then, during the dance, have no control at all over the unruly crowd. You may still fairly be held to account with a sanction (e.g., be fired), given that you could have avoided the sanction by not agreeing to take on the job in the first place (Watson, 2004: 276).

[4] Bad actions that are *justified* (and so also excused) would require a different grounding, however.

Nevertheless, recent work casts doubt on whether voluntarism's exclusive focus on actions can capture the entire realm of responsibility, for it looks as if we may also be responsible for a wide range of *attitudes*. These include, most specifically, bad attitudes like malice, contempt, a "lack of hearty concern" for the welfare of others, corrupt beliefs, and wrong desires (Adams 1985: 4); more generally, "beliefs, intentions, hopes, fears, ... admiration, respect..., indignation..., the view that fame is worth seeking" (Scanlon 1998: 20), guilt, envy, shame, and jealousy (Smith, 2005: 254, 258); and, most generally, certain things people notice, neglect, and forget, certain thoughts and impulses, and certain spontaneous reactions (Smith, 2005: 241). These attitudes for which we are responsible are in contrast to those psychic states for which we are not: "mere feelings such as hunger...or tiredness or distraction" (Scanlon 1998: 20); "thirst, and the desire to eliminate bodily wastes" (Smith, 2005: 248); and "[n]onintentional mental states, such as physical pains, sensations, and physiological conditions..." as well as some "random thoughts and images" (Smith, 2005: 257, 260).

The difference between the attitudes for which we're taken to be responsible and the psychic states for which we are not is that it is only the former that we may reasonably be asked to defend, to *answer* for (Smith, 2005: 251). These attitudes for which we are (in principle) answerable, therefore, must be *judgment-sensitive*, that is, "attitudes that an ideally rational person would come to have whenever that person judged there to be sufficient reasons for them and that would, in an ideally rational person, 'extinguish' when that person judged them not to be supported by reasons of the appropriate kind" (Scanlon 1998: 20). So to the extent that my rational evaluative mechanisms are functioning properly, my fear of the snake you are waving around, say, should disappear once I recognize that it is rubber, and so not worth fearing. This is not to say that my fear *will* disappear. Rather, as Smith points out, "The 'should' in question here is the 'should' of rationality and, therefore, marks a normative ideal which our actual attitudes may not always meet" (Smith, 2005: 253). When my attitudes do not in fact reflect my evaluative judgments (or when I hold contradictory evaluative judgments), I am irrational, and in being irrational I am blameworthy, that is, criticizable.[5]

[5] *How* criticizable I am, however, will be a function of other factors, including how bad the attitude is and perhaps my history in coming to have it. See Smith, 2005: 267-69.

How, though, does appeal to judgment-sensitivity explain an agent's responsibility for *forgetting* something, for failing to *notice* something, or for lacking "a hearty concern" for the welfare of others, none of which are actually attitudes? Evaluative judgment is thought to govern a "complicated set of dispositions to think and react in specified ways" (Scanlon 1998: 21). Suppose that I judge that my friend is worth caring about. Ostensibly, evaluative judgment rationally entails all sorts of behavioral and attitudinal dispositions: If I truly care about him, I ought to check in on him regularly, to desire his well-being, to be available to take him to the airport or help him move, to notice when he's slighted by others, to call him up on his birthday, and to be joyful at his triumphs and saddened by his hurts and losses. But this means that when I *fail* to have these dispositions activated – for example, by forgetting his birthday, not noticing slights toward him, or not getting emotionally disturbed when something bad happens to him – then I am responsible for these failures insofar as my *actual* judgments about his value and significance are reflected in them. *If I really cared about him* (or cared more), then I *would have* remembered his birthday, noticed the slight, or have been sad alongside him. So I am responsible – answerable and, further, criticizable – for forgettings, noticings, and so forth, because they seem to reveal that I don't care as much as I claim I do, and if that is true then I have not actually judged my friend to be as significant or valuable (worth caring about) as we both, perhaps, had thought. This account has been most thoroughly developed by Angela Smith, as what she calls the *rational relations view* (Smith, 2005).

Why think we are responsible for judgment-sensitive attitudes? There are two arguments. The first, as I have already noted, simply applies a general conceptual commitment to responsibility as, most fundamentally, about answerability (see, e.g., Duff, 2007, 2009, 2010; Smith, 2015). As Scanlon puts it, "Because 'being responsible' is mainly a matter of the appropriateness of demanding reasons, it is enough that the attitude in question be a judgment-sensitive one – that is, one that either directly reflects the agent's judgment or is supposed to be governed by it" (Scanlon 1998: 22). Or as Duff puts it, the first stage of holding people responsible for something occurs when "we attribute [it] to the person as its author, and request (or demand) that she answer for it" (Duff, 2009: 980).

The second argument is simple: Our practices reveal that we *do* in fact attribute such attitudes to people for purposes of responsibility, that we do de-

mand answers for why they had them and then criticize them in light of their unacceptable answers. That is, our blaming patterns and practices target people not only for their actions but also for their attitudes. If a theory of responsibility must account for the data of our blaming practices (as many think), then it will have to range over more than intentional actions (Smith, 2005: 236-7, 240, 241; see also the discussion in McKenna, 2012: 187-195). (Keep this conditional in mind later!)

If these authors are right that we can be responsible for our attitudes, and voluntarism maintains that we can be responsible only for those things over which we are able to exercise volitional control, then voluntarism is false. Most attitudes are not governable by volitional control.[6] I cannot, as acts of direct will, consciously choose to remember that it is my friend's birthday or to cease being amused by a racist joke. If we are nevertheless to include such items as things for which we are responsible, and if responsibility requires control, then it looks like we need a non-voluntarist conception of control (Smith, 2005: 265; Hieronymi, 2006, 2014; McHugh, 2017; Portmore, 2019).[7]

Call it *evaluative control* (Hieronymi, 2006, 2009, 2014).[8] On Smith's version of this view, one is responsible for an attitude if and only if it has a "rational relation" to one's judgments about the worth of reasons: "[I]t must be the kind of thing that either directly reflects, or is supposed to be governed by, [one's] evaluative judgments" (Smith, 2005: 256; see also Smith, 2012). Consequently, were I to change my judgments about the worth of some reasons,

[6] This is essentially the main argument of Smith (2005).

[7] There are other options, of course, including rejecting responsibility for attitudes (given that they're not governed by voluntary control; see Wallace, 1994: 131-32; Sidwick, 1981: 60-61; Taylor 1970: 241-52); showing how we can actually trace the attitudes for which we are responsible to some opportunity for voluntary control (akin to Levy 2005 and 2014; Fischer and Tognazzini 2009); or denying that control is necessary for responsibility (e.g., Sher 2006; Scanlon 2008: 193-98). I'm interested here only in how expanding the arsenal to include non-voluntary control opens a wedge to including one more type of control. I will address some of these options at the end, however.

[8] I am here borrowing Hieronymi's label, as I think it comes closest to describing the answerability form of responsibility most closely associated with it. Several other labels are available, though. Smith uses the label "rational control" (Smith, 2005: 265), as does Portmore (2019: 27). Scanlon occasionally uses "reflective control" to describe a notion in this neighborhood (Scanlon, 2008: 194). McHugh (2017) uses "attitudinal control," which somewhat oddly refers directly to those things being governed, and not the process by which they are governed (would its contrast be "actional control"?).

my attitudes ought to – if I am rational – fall in line. This counts as a kind of control, it seems, because among the familiar connotations of "our having control over X" is the notion of "X's being *up to us*," and evaluatively-controlled attitudes are indeed, in the relevant answerability sense, up to us, "that is, they depend on *our* judgment as to whether appropriate reasons are present" (Scanlon 1998: 22; emphasis mine).

Some defenders of this view have argued that attitudes are the only things for which we are directly responsible, as actions merely execute our choices or intentions, and so we're only derivatively responsible for them (Scanlon 1998: 21; see also McHugh and Way, 2022). We need not be so dramatically restrictive, however, and anyway there does seem to remain a gap between choices/intentions and actions, most clearly brought out in cases of weakness of will ("I chose and intended to stick to my diet tonight, but I found myself reaching for the cake regardless."). And this gap seems bridgeable only by a different form of control than what generates and governs those choice-associated intentions (or other attitudes). It is not as if I – even if I am ideally rational – can execute my intention to perform some action merely by reaffirming my judgment that doing that thing would be good. Rather, I must exercise *volition*. I think it less controversial, then, to maintain for purposes of this paper a kind of pluralism with respect to the control condition. Control must come in two flavors: Volitional control, we may say, governs the *actions* for which agents are responsible, and evaluative control governs the *attitudes* for which agents are responsible.

However, once we allow evaluative control as a second type of control, we have introduced a wedge for admitting a third type of control into the mix, for there is an additional psychic stance for which we are often held responsible, even though it is not governed by either volitional or evaluative control.

3. Regard and Disregard

The fundamental target at the heart of interpersonal accountability is regard – or disregard – for others. To properly regard others is, roughly, to take them sufficiently seriously (see Shoemaker, 2015: Ch. 3).[9] We obviously demand re-

[9] The range of "others" could also include one's future self. See Shoemaker (2020) for an account of how this works that dovetails with the present paper.

gard from one another (see, e.g., Strawson 1962/2003; Wallace 1994; Watson, 2004: 219-259; Darwall, 2006; Shoemaker, 2007; and McKenna, 2012: Ch. 3). Violations of the demand for regard conjure up blaming responses in the emotional neighborhood of anger.

Everything I've just said is rather vague. But regard is a slippery notion. This is in part because there are many ways in which we can be said to take people sufficiently seriously. We can do so by treating them properly, or we can do so by having the proper sorts of attitudes toward them. But we can also do so, I will now show, by *perceiving* them properly.

The best way to start getting a more precise handle on regard is by thinking about clear and familiar cases of *disregard*. Active, malicious disregard – aiming to hurt someone, say – is actually kind of rare, which is why it shows up on the news. Much more typical are cases in which we simply fail to sufficiently regard those our actions and attitudes negatively affect. Perhaps I'm supposed to pick up my daughter after school, but I'm having such a good time golfing with my buddies that I don't notice how the time has gotten away from me, and I leave her stranded for a couple of hours (drawn from McKenna, 2012: 59). Or perhaps I'm throwing a party, and I let the loud music play on till well past midnight, preventing my next-door neighbor from being able to sleep. Or perhaps I don't tell my roommate that the dream date phone call she'd been dying to receive came in last week, because I'm hoping that I have a romantic shot with her instead.

These are all cases in which I wrongly treat my own interests as having greater weight than someone else's. There are several subtly different ways to make this mistake, though. Take the party case, where someone else at the party has turned up the music, and I ought to turn it down but don't. There are three different explanations for my failure to turn it down: (1) I might know full well that at that level the music is going to keep my neighbor awake, but because I dislike her, I think, "Screw it!" and do nothing; (2) I might know full well that at that level the music is going to keep my neighbor awake, but because the music is just so thrilling to hear that loud, and the party is just so much fun, my neighbor's being awake starts to seem way less important to me, so I do nothing; or (3) I might be having so much fun that I simply fail to register the fact that the music will keep my neighbor awake, so I do nothing. Each of these three possibilities involves a different sort of disregard for my neighbor, insofar as each consists in a different sort of mistaken attitude toward the fact that keeping the music so loud will keep my neighbor awake: In (1), I treat that fact as a reason to

leave the music's level where it is; in (2) I see that fact as a reason to turn down the music but I discount its relative weight; and in (3) that fact doesn't even appear to me as a putative reason to turn down the music in the first place.

There are similar forms of disregard in our emotional reactions, for which we also have reasons (of *fit*), reasons toward which we may also have mistaken (disregarding) attitudes.[10] Perhaps you are watching a female employee being publicly berated – quite unfairly – by her boss, yet you feel nothing. Perhaps you even feel amusement. This is disregard, but it may have the same three sorts of explanations: (1) You hate that employee, so you treat the fact of her public humiliation as a reason for enjoyment; (2) you register the fact of her public humiliation as a reason to feel sympathy or concern, but if you do actually feel such emotions, your inevitable expressions of them may draw the attention of the boss who would then turn on you, so you mind your own business and wind up feeling nothing; or (3) because the employee's humiliated face seems to have a cartoonish look to it, you're amused, and you don't even register the fact that she's being publicly humiliated.

In both examples, the first two types of disregard reflect poor evaluative judgment. These types have as their source mistaken judgments about, respectively, the status or the weight of some actual reason (e.g., the fact that it will keep the neighbor awake, the fact that the employee is being publicly humiliated). In the first type of mistake, the agent incorrectly judges some fact to be a reason for the wrong action or attitude. In the second type of mistake, the agent correctly judges that fact to be a reason for the right action or attitude but then incorrectly judges that reason to have the wrong weight, at least relative to that agent's own selfish interests.

The third type of mistake, in both instances, does not consist in a mistaken judgment about reasons. To make the first two types of mistakes about reasons, the relevant facts have to at least *appear* to you to be reasons. But some facts just don't. And some facts that *do* appear to be reasons actually aren't, but these putative reasons may continue to appear to you as reasons even after you judge that they aren't. As Scanlon notes about them, "Such 'seemings' arise independently of our judgment, and they can, unfortunately, persist in the face of it" (Scanlon 1998: 65). But sometimes these very *seemings* are the source of a

[10] On reasons of fit for a certain range of emotions, see D'Arms and Jacobson (2000; forthcoming).

third type of disregard. Sometimes facts appear to you in what I will call a "reasonish" light when they shouldn't, as when the employee's humiliated face seems like it generates a comic reason for amusement. At other times, facts that should appear reasonish don't, as when my neighbor's being kept awake by my pounding music simply doesn't register in my deliberations at all.

This third variation I will label *reasonish regard*. When various facts about others and their interests properly appear to me as putative reasons, I have reasonish regard for them; when they don't, I don't. Reasonish regard is sensitive to neither volition nor evaluation. Yet its failure is something that can aptly rouse blaming anger.

Or so I shall argue over the next two sections. I'll take the latter point first.

4. Blameworthy Failures of Reasonish Regard

We blame people for mere perceptual failures in all sorts of domains. Start with straightforward visual perception, where morality and disregard are nowhere in sight. Suppose Dick has been raised in a duck dynasty, in a remote bayou where everyone obsesses about ducks, talks only about ducks, hunts ducks, and looks at pictures only of ducks. Once grown, he leaves Louisiana, and he takes a philosophy class in which he is presented with the famous "duck-rabbit" image (*Figure 1*). He only sees it as a duck.[11] No matter how much he scrunches up his will or rehearses how good it would be to see the image as a rabbit (his grade will be affected by his failure, say), neither attempt has any effect on what it is that he sees, because both are irrelevant to generating visual perception. He of course will have good reason to *attend* to various features of the image, and he can presumably voluntarily attend in the requisite way, but whether he *sees* the rabbit is not governed by either.[12]

[11] This sort of thing is clearly possible, even though it doesn't happen very often in response to what are known as bilateral images such as the duck-rabbit or the Necker cube. Ordinary perceivers typically switch back and forth between seeing the two images, but they also tend to have one dominant perception, seeing the image more often as one or the other. I take Dick to be a limiting case of this phenomenon. Such "perceptual biases" have been found often to be a function of stability in viewing conditions rather than antecedent perceptual experience (see Klink, 2008), but the tendency to bias strength itself may also be a function of prior probability knowledge or (in)sensitivities to various kinds of information about consequences (Zhang et al., 2017).

[12] This is akin to what Audi (2013: 10) calls "attributive perception."

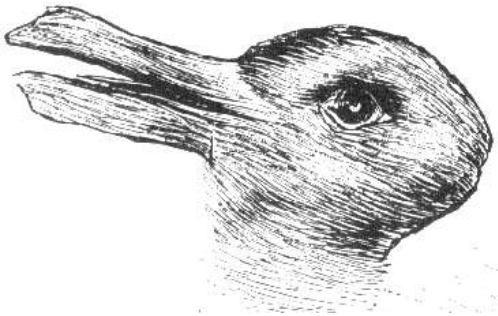

Figure 1

Having Dick around could be frustrating and irritating, and could lead some to criticize him: "What's *wrong* with you? Can't you see that bill as the rabbit ears, and can't you see the eyes as looking to the right?" This frustrated criticism presupposes that Dick *can* somehow see it, that he has the capacity to do so. And there's no reason to think that his childhood experiences eliminated that capacity. So the criticism may make sense.

It's also the sort of criticism teachers of logic occasionally launch at students who can't see the logical force of *modus ponens*. They obviously cannot simply exercise their will and see it, nor can they simply see it by judging that it would be *worth* seeing it (a good grade *is* worth it!). The demanded perception occurs, when it does, independently of both volition and evaluative judgments. It's as if an obscuring veil has suddenly been lifted.

Both of these cases allow for voluntary control over *attending* to a subset of perceived properties, and they both allow for evaluative control over *intending* to attend to those properties. But there remains a gap between their attending to this subset of properties and their coming to *see* the additional salient property in each case, respectively, "rabbitness," and "logical force." And when they fail to bridge this gap, they may well be criticized, under the presupposition that there is *flawed agency* at work: they are capable of seeing it, but they don't. If this is correct, it could make such criticism apt, assuming, that is, they have some sort of *control* over the exercise of their capacities.

This last point is likely controversial. We may feel in such cases that nothing much turns on successful perception, or at least that we can't reasonably *demand* such perceptions from people. Consequently, when there are failures, they aren't really criticizable, and those who criticize do so unreasonably.

Let us then consider cases in normative domains where there are indeed demands for perception, and where failures are aptly criticizable, that is, blamable. The best nonmoral example comes from humor. Very often, apt amusement depends on perceiving various properties of objects or states of affairs, as well as perceiving how they interact, namely, in the amusing way. But there are those who just don't see slapstick, dark humor, or clever insults *as* funny (as meriting amusement). What they see are the same facts as the rest of us – for example, various incongruous properties, to draw from a leading theory of humor – but they don't see those facts as reasonish (as putative reasons) for amusement; indeed, they may mistakenly see those facts exclusively as reasons for anger or offense instead.[13]

For example, suppose that close friends express affection by exchanging teasing insults. A stranger may walk in on this and mistakenly perceive the properties of these "insulting" exchanges – a combination of wordplay, tone, timing, and intent (among many other properties, surely) – as exclusively constituting instead *a slight*, and so calling for anger. But the combination constitutes only a faux and funny (non-slighting) insult, and so offers not only putative but *actual* reasons (exclusively) for amusement. The person who fails to see the facts in this way may be criticizable for having a poor sense of humor (see Shoemaker, 2018).[14]

I say *may* be criticizable. One might have a poor sense of humor in a couple of different ways. One might simply be incapacitated for recognizing some reasons for amusement. This is akin to those who have poor senses of sight or hearing: it's a true description of them that they don't see or hear well, but these failures aren't down to them in a way that grounds criticizability. That would require a functional capacity for seeing or hearing that they simply don't exercise well. So too with criticizable senses of humor: They require a functional capacity for seeing and responding to reasons for amusement that is poorly exercised. People like this we label *prigs*. They fail to see various funny-making

[13] The "exclusively" matters, as some facts can serve double-duty, counseling in favor of both amusement *and* anger, as in some cruel jokes, where the cruelty is a crucial part of what makes the joke funny. Or so I argue in my unpublished monograph Wisecracks.

[14] This is different than what D'Arms and Jacobson (2000) call the "moralistic fallacy," which takes moral considerations to determine not only what one ought not to be amused by but also, fallaciously, what's rendered not funny thereby. In the case I've given, there aren't even any moral reasons in play.

facts as even putative reasons for amusement, perhaps due to the fact that they are hyper-sensitive to reasons for moral outrage, or perhaps because they see the world as too tragic a place for amusement to have any place (see Debbie Downer).

Criticism is apt for those who poorly exercise functional aesthetic perceptual sensibilities generally, those who simply don't see various aesthetic facts in a reasonish way, as being at least putative reasons for amusement, admiration, awe, or other aesthetic emotions. And so too there are criticizable failures of perception throughout our *morally*-charged interpersonal interactions. Here are two examples:

Black Lives Matter: A defender of the *Black Lives Matter* movement talks with a defender of the *Blue Lives Matter* movement about white police officers shooting unarmed black people around the USA. The Blue Lives Matter defender is quite upset with members of the Black Lives Matter movement generally, in particular over the facts that they are so angry, that they are protesting and lashing out at "those who protect us," and that the Black Lives Matter protests tend to incite violence against the police and private property. The Black Lives Matter defender then says to the Blue Lives Matter defender: "But just try to see things from my perspective as a black man. Suppose you had been raised in a community where you and people like you were harassed daily by the police, followed around in stores, and viewed as 'suspicious' your whole life. Suppose that stories about cops killing members of your community were common knowledge for years, and that you were fearful, both for yourself and for your children, every time you heard a police siren, not knowing if they were coming for any of you and what they would do to you when they found you. Wouldn't you be angry too, having to live life in this way?" The Blue Lives Matter defender pauses and, in realizing it would be good to do so for the sake of honest dialogue, actually considers it: What *would* it be like to be black man? Nevertheless, as she takes up his perspective, all she continues to see are the reasons she would have (*as herself* in his shoes) to obey and support the cops, and to be grateful when seeing them arriving on a scene, insofar as they are the protectors of the innocent. Consequently, she simply repeats, "The police are here to protect us, and their lives matter just as much as black lives do."

Don't Fix It: Ayisha comes home from a work meeting visibly upset. Her spouse Rahul asks what is wrong, and she tells him: Her boss didn't consult the opinion of any women at the meeting, asking only the men what they thought. Rahul calmly and efficiently starts counseling Ayisha on what she should do in response: She needs to send her boss an angry email, and then contact his

> superior, and then rally the other women to send him a list of demands, and so forth. Ayisha stops Rahul in his tracks, saying, "I just want you to *understand and appreciate* what it was like for me, sitting in that meeting, not being acknowledged. I don't want you to 'solve the problem' or tell me what to do; I *know* what to do." So Rahul imagines *himself* at her meeting, and feels himself getting all riled up with righteous anger at the injustice of being ignored, and still only sees reasons to right the wrong, to resolve the problem. Thus, in response to Ayisha's continuing demand that he simply understand and appreciate what it was like for her, Rahul responds: "But how will *that* solve anything?"

In **Black Lives Matter**, The Blue Lives Matter defender fails to see what the Black Lives Matter defender is demanding that she see. It's not a demand simply to see what life would be like *for her* while occupying a black man's location in the world; it's rather a demand for her to see what it's like to be *him* in that location and so to experience the world as *he* sees it (with his ends and emotions), and so come to perceive *his* powerful reasons for anger as at least putative reasons for *her* to feel angry at the injustice. The defender of Blue Lives Matter, let us stipulate, does genuinely believe that *all* lives matter (even equally). But she cannot as an act of volition come to see the distinctive reasons black people have for anger, nor could her evaluative judgment that it would be good to see such reasons generate her actually seeing them.[15] Nevertheless, this failure of perception could well be criticizable: In being capable of seeing these facts as putative reasons for her, but in failing to see them as such nevertheless, the Blue Lives Matter defender has a blamably poor sense of justice.

In **Don't Fix It**, Ayisha is clearly making an interpersonal demand that Rahul is failing to meet, namely that he simply come to understand and emotionally appreciate what things were like for her at that meeting. She is demanding this simply so that they can share, as much as possible, in this traumatic experience (call this *emotional communion*). That is quite often taken to be a legitimate demand in a close loving relationship like theirs. Were Rahul to meet this demand, presumably, the facts he would come to see in a reasonish light would not (just) be facts about her problem and how to solve it, but rather facts about her need to be understood and acknowledged. He should also come to have a greater emotional appreciation than he does of the distinctive sort of hurt and anger she feels as a woman being treated this way yet again in a man's world. This

[15] I recognize that many will balk at this claim. I deal explicitly with this balking in a later section.

is not necessarily the righteous indignation of the unjustly oppressed; it instead may be a hurt and anger tinged with shame, at having her gender pointed out in a disparaging way yet again (cf., Webster, 2021). Rahul cannot accomplish these things by either volition or judgment. Nevertheless, his failure to accomplish them is criticizable: Assuming he has a functional capacity to understand and appreciate things from her perspective, he has a blamably flawed relationship sensibility.

These are both failures of reasonish regard, failures to perceive certain facts about others in a reasonish way, a demand for which we often take to be fundamentally necessary for taking each other sufficiently seriously. They also both seem to be aptly criticizable failures in a way that implicates *control*: Insofar as they are failures to live up to a certain sort of legitimate demand (to see certain facts in a reasonish light), and insofar as legitimate demands are taken to presuppose the capacity to adhere to them, legitimate demands presuppose some sort of control. But neither of the failures just discussed seems to implicate voluntary *or* evaluative control. So some other form of control is being presupposed. What might that be?

5. Empathic Control

Reasonish regard is a matter of how various facts appear to agents. This is a passive perceptual stance. It is thus itself neither an action nor an attitude. It is instead a state of *empathy*.

Unfortunately, the psychological and philosophical literature has made a hash out of empathy, so much so that there's just no shared understanding of it to be had (cf. Nichols, 2004: 30-31). We will have to make do with my stipulation of it as applied to a very familiar psychological phenomenon (a phenomenon that I do think is best captured by the concept of "empathy," even though I won't fight for it here).

Empathy is often supposed, in the psychological literature, to consist in some kind of *perspective-taking* (called "cognitive empathy"). But to the extent that empathy and sympathy should have some kind of conceptual kinship, and sympathy is a concerned or compassionate *response* to the plight of others

(for their sakes), then empathy is better construed, I think, as a response as well, specifically a response to perspective-taking of a particularly robust kind.[16]

Robust perspective-taking requires viewing and trying on other people's interests and ends in the way that they do, namely, as *worthy* of attention and pursuit, and so as things to which they are disposed to respond emotionally when damaged or promoted. It's more than merely taking up someone's perspective in order to identify what their interests and ends are (through the prism of your own ends and interests); psychopaths are capable of that much, and they use that ability to exploit other people. Rather, it has to involve appreciating the reasons that person has for pursuing those ends and interests, and then being emotionally susceptible to feeling threats to and promotion of those values as that person does, as *mattering*.

Taking up someone's perspective in this robust way will tend you toward two kinds of responses. First, when you return to your own deliberative perspective, you are liable to persist in seeing their ends and interests in a reasonish way, as giving you at least putative reasons to take seriously those interests and ends in your *own* attitudes and actions. Second, when you are emotionally vulnerable in taking up their perspective, you will typically come to feel, vicariously, *as they do* about promotions of or setbacks to their interests and ends when they occur (Denham, 2017). These two sorts of responses are what I am labeling *empathic*.

Let's then connect this up with the idea of control. Minimally, "agential control" refers to "the manner in which a particular piece of behavior is connected to, controlled by, or an 'outflowing of' the agent" (O'Connor, 2000: 23; see also, Ginet 1990, and Fischer and Ravizza 1998). Of course, if we have evaluative control over our *attitudes*, as I've been allowing, then we need to expand the definition to include more than behavior. "Control" is most often construed as implicating a causal mechanism, but some theorists think agential control may be noncausal (as in Ginet 1990), so I don't want to beg any questions with the minimal definition. In any event, if reasonish regard just consists in empathy,

[16] There is actually quite a bit of agreement on the nature of sympathy, oddly enough. See Darwall 2002: 50-53 for a nice discussion. But again, I should stress, I don't care if the phenomenon I describe *really is* empathy or not; all I care about is the phenomenon itself and its role in our interpersonal lives. What I describe has some kinship, however, with what those in the psychological literature call emotional (or affective) empathy. See Adam Smith (2006).

and empathizing is the relevant "outflowing of" a functional agent's robust perspective-taking, then responsibility for reasonish (dis)regard requires what we can call *empathic control*. Thus, in order to be the proper target of blame for reasonish (dis)regard, an agent must have (and be able to exercise) a capacity for empathy.

There is a natural way of thinking about empathic control that makes it perfectly continuous with voluntary and evaluative control, and it comes from adopting the language often adopted by free will and action theorists, namely, the language of *mechanisms*. Mechanisms are not to be thought of as reified objects in the brain; rather, this terminology is just shorthand, in the case of action, for "the *process* that leads to the action, or the '*way* the action comes about'" (Fischer and Ravizza 1998: 38; emphasis added). So to say that one has "voluntary control" over some action is just to say that one has a normal volitional mechanism wherein the action is a function of a deliberate choice or intention. Contrast cases – which generate excuses from responsibility – include hypnotized, manipulated, or brainwashed actions, which are instead a function of some *ab*normal mechanism. It is easy to appeal to mechanism talk for evaluative control as well. To say that one has "evaluative control" over some attitude is just to say that one has a normal rational mechanism wherein the attitude is a function of an evaluative judgment. Again, hypnosis, manipulation, and brainwashing are contrast (excusing) cases, but so too are various phobias, manias, and compulsions.

Empathic control could be constructed analogously. To say that one has "empathic control" over whether one has reasonish regard for someone is just to say that one has a normal empathic mechanism wherein the demanded perceptual stance is caused or brought about by robust perspective-taking with that person. So were I to think about how loud and irritating my music must be from the perspective of my neighbor, how she values work and is livid over this din, I should come to perceive those facts from my own perspective as at least putative reasons to turn it down. If I were a white supporter of Blue Lives Matter, but I were to fully take up the perspective of a black man in America and see his ends and interests as he does, I should come to perceive various facts about those ends and interests as reasonish for me, and to see as reasonish my feeling vicarious anger at his mistreatment as well. And if I were to fully envision what it must have felt like for you, my partner, to have been spoken over and ignored by your boss, I will see as reasonish my feeling upset and ashamed alongside you in the

way you do, and perhaps I will also see, from my own perspective, at least putative reasons *not* to try and solve your problem. However, if reasonish (dis)regard were the product of some abnormal mechanism in any of these cases, I would not be responsible or criticizable for it. Perhaps, for instance, I am a nurse whose empathic mechanism has been completely exhausted by yet another day of dealing with too many suffering patients. When you ask me to feel what it was like for you at work, I may be excused from any criticism if I can't summon any more empathy today. Here my stress and emotional exhaustion may swamp my ability to respond as you'd like me to.

The mechanism-based approach generates a way to respond to one objection I'm sure many are having right now, namely, while the first two types of control seem very agential, the last one doesn't. After all, when the student finally has the "aha!" moment when recognizing the logical force of modus ponens, that revelation seems to occur quite independently of her agency. So too it may be that when we do come to robustly empathize with our fellows, it also occurs – aha! – independently of our agency. Indeed, I already described this process earlier as "passive."[17]

But again, following Fischer and Ravizza, our agential mechanisms (in their case, for reasons-receptivity) "must be at least minimally grounded in reality" (Fischer and Ravizza, 1998: 90). These mechanisms have to be receptive to the way the world actually is, so as to enable us to respond to it in an intelligible and coherent way (our pattern of response has to be "understandable by some appropriate external observer" [Fischer and Ravizza 1998: 90]). Again, they are talking about reasons, whereas I am talking about the appreciation of the world in a way that enables one to see facts as reason*ish*. Nevertheless, our being properly attuned to the real normative world, receptive to the special status of various moral facts, is *agential*, even if not fully under our volitional or rational ken. Indeed, one might say, seeing various facts under the proper (moral) light is the most fundamental of our agential capacities, as it's what in fact seems necessary for going on to make evaluative judgments about those reasons and subsequently acting on them.[18]

[17] I am quite grateful to an anonymous referee, both for raising this objection and for suggesting the very helpful response that follows.

[18] One question that I lack space or ability to delve further into here is whether the mechanism-based understanding of control is the only way to explain the nature of empathic control that is continuous with volitional and evaluational versions, or if it's also possible to explicate it by appeal

6. Reasonish Without Reason

It is, I think, fairly obvious that getting to a state of reasonish regard is not a matter of voluntary control. One cannot come to see emotional and moral facts as reasonish via an act of will. But I have also been merely asserting that neither is it a matter of evaluative control, of *judgment*. And here many may balk. Why wouldn't my cases instead fall under the rubric of *noticings* or *neglectings*, which are indeed, as Angela Smith (2005) has compellingly argued, sensitive to evaluative judgments. Why aren't they things that *can* be up to us in the way necessary for us to be answerable for them? As she puts it:

> [W]e do take people to be open to criticism for failing to notice morally (or interpersonally) salient features of the situations in which they act. A person who systematically fails to notice features of situations which bear on the welfare of others, for example, does seem to be guilty of a normative failure of some kind. Philosophers sometimes refer to this as a failure of "moral perception," which itself suggests that our perceptions are attributable to us for purposes of moral assessment. But these are usually cases of what might be called "seeing under an aspect," where what is in question is not a person's visual perception per se, but the significance (or lack of significance) that she attaches to what she perceives. A morally insensitive person may, in a literal perceptual sense, "see" exactly the same thing as a morally sensitive person – for example, that a person is standing on a crowded subway with two very full grocery bags. What differs is that the morally sensitive person sees this person *as* uncomfortable and in need of a place to sit down, while the morally insensitive person does not (Smith, 2005: 259; example from Blum 1994: 32-33).

Thus, while literal visual perception is not (typically) sensitive to our evaluative judgments, the *significance* of what we literally see can be. And it might even be that, to the extent that certain evaluations can directly shape what we *literally* perceive, those literal perceptions may be attributable to us – to our evaluative judgments – as well (Smith, 2005: 260).

These remarks generate the most obvious, and most powerful, objection to my view. If our fellow moral agents matter to us as they should, then the

to *agent-level* theories of control, such as in McKenna (2013). My hunch is that it might have more difficulty on the latter sort of theory, but again, I'm unsure. And as my aim in this paper is simply to sketch the kind of control that seems to be required by some aspects of our practices of responsibility attribution, I can safely leave this work in the hands of those who do deem it necessary.

reasonish regard I have proposed should be rationally connected to that evaluative judgment. In other words, why wouldn't it be a rational expectation that those who truly judge others to be worthy of sufficient regard will *thereby* come to see facts about their interests, ends, and emotional states as at least putative reasons from within their own deliberative and emotional frameworks?

Let's start with some unity: Smith and I agree that reasonish regard or disregard is something for which agents may aptly be held responsible. What we disagree about is whether this empathic state is rationally related to evaluative judgment. There are several problems with her claim that it is.

First, the notion of evaluative judgment is quite ambiguous, so much so that one may well worry that it captures everything Smith wants it to capture without convenient and subtle shifts of meaning. While it originally sounds as if the notion is highly intellectual, conscious, and active, a matter of intentionally putting one's stamp of approval on various propositions, it later gravitates into a more all-purpose *mattering*: "[T]he judgments I am concerned with are not necessarily consciously held propositional beliefs, but rather tendencies to regard certain things as having evaluative significance" (Smith, 2005: 251). But then that includes not just what we would ordinarily call "judgments" but also "things we care about or regard as important or significant," and they "are often things we discover about ourselves through our responses to questions or to situations" (Smith, 2005: 252).

Once Smith expands the notion of judgments to include cares and self-discoveries, however, she runs into two problems. First, the things we care about are sometimes beyond the purview of reason or justification, and sometimes, second, what we "discover" about ourselves through questioning is a matter of simple confabulation. So in including these phenomena, she has simply gone beyond the boundaries of true *answerability*.

I have made the first point repeatedly elsewhere (see, e.g., Shoemaker, 2011). Cares are emotional dispositions, and sometimes they manifest in attitudes for which the carer is the appropriate target of responsibility responses such as admiration, disdain, or contempt, even though they seem to arise independently of any evaluation or judgment. These include desires for revenge, as when I care deeply about ensuring someone's destruction, despite my clear-eyed view of just how terrible it will be; volitional necessity, as when I can't bring myself to put a limping rat that's found its way into my home out of its misery, and which I know will *increase* its misery (see Frankfurt, 1988: 80-94 for the

label; see also Shoemaker, 2015); moral dumbfounding, as when I have a powerful aversion to eating human flesh, even though I'll starve otherwise (see Haidt, 2001); and cases of parental love, which, in an unconditional form, may persist despite one's child being a serial killer unworthy of care. The question, "Why do you want revenge so much when you know it's going to destroy everything you hold dear?" or "Why do you still love the guy who hurt you and who you know full well is worthless?" *have no answers grounded in evaluative judgment*. Instead, we may find ourselves saying, "I don't know, I just do!"[19]

Nevertheless, even though agents aren't answerable for being in these states, we still view their attitudes as *attributable* to them for purposes of responsibility: The attitudes definitely belong to them, where what that means is that they are manifestations of *who they are as agents*.[20] We also take them to be criticizable for those attitudes, insofar as we respond with admiration, disdain, and contempt to these aspects of agency. In other words, in our actual moral practices (practices which Smith takes as her guide; see Smith, 2005: 263), we don't take answerability to be coextensive with attributability (or basic responsibility most generally; see the coextension view defended in Smith, 2012; for a reply, see Shoemaker, 2015: Ch. 2). The latter ranges more widely than the former. People can fail in ways that ground various emotional responsibility responses, but without being answerable for those failures in ways Smith would have us believe. These responses also ground how we view and interact with such people generally. Those who tend toward vengeful anger that's insensitive to their evaluative judgments tend to be given wide berth.

[19] This response is compatible with there being an *explanatory* answer: perhaps it's due to the agent's tortured relationship to an emotionally unavailable parent. But that's not the sort of *evaluative* reason necessary for Smithian answerability. I'm grateful to an anonymous referee for this note.

[20] What if they don't view themselves as attributable? This question (raised by an anonymous referee) gets us into troubled waters. I here defer to my previous work on the topic (Shoemaker, 2015: Part Two), which does attempt to put forward an objective account of attributability. It is favored, in small part, by considering differing first- and third-personal judgments of attributability in cases of clinical depression (where the depressed person tends to attribute negative traits to herself, whereas third parties overwhelmingly tend not to), as well as other instances (like OCD), where it looks like the first-person perspective is clouded in significant ways. But third-person judgments can be clouded too, I realize. Again, this is a large topic that I can't explore further here.

Similarly, I think, we would likely criticize the morally insensitive person in Smith's case: He *ought* to see the person with the shopping bags as uncomfortable and in need of a place to sit down. Smith's claim is that this demand is an answerability demand, and his *seeing as* should have a rational connection to his evaluative judgments. But notice that being able to *see as* in this way is not enough for moral sensitivity, given that a psychopath has the capacity to do that much. The psychopath sees the person's discomfort as a reason to be amused, or perhaps as a reason to remove all the chairs on the subway platform. No, what we want from the morally sensitive person is that he will see the person's discomfort as a reason to feel vicarious discomfort and as at least a putative reason to *resolve* that person's discomfort from within his own deliberative framework. But those perceptual states just aren't sensitive to evaluative judgment (and they certainly aren't subject to his voluntary control); instead, they could only be a function of empathic control.[21]

Let's move on to the second general worry about Smith's claim that the relevant perceptual states are a function of evaluative judgment. There is a mountain of evidence that we are thoroughgoing rationalizers and confabulators. We are easily able to – and constantly do – respond to answerability demands with "answers" that, while seeming to manifest discoveries of evaluative judgments we hold, just *aren't*. That is to say, we make up the answers! As John Doris puts it, "[P]eople readily interact in the 'space of reasons,' even when this space fails to overlap with the 'space of motives' where the causal and psychological origins of behavior are found" (Doris, 2015: 141).

There is enormous social pressure, when we are presented with the answerability demand, to *say something*; *anything*; and that's what we tend to do. But saying something doesn't imply that what we say or even think in response to that question has anything at all to do with the grounds of our actual emotional responses or the way they tend to motivate us (this point is also a major theme of Haidt, 2001). Our motives are often instead a function of our emotional sensibilities, which themselves may well appraise their objects independently of all evaluative judgment.[22] What counts as the so-called "rational relation" between

[21] What all of this also means, interestingly, is that demandability doesn't presuppose answerability either.

[22] I actually take this to be a powerful truth about emotions. See, e.g., D'Arms and Jacobson, 2003, Scarantino, 2014 for lots of good arguments in favor of it. Developing and defending this view of emotions here would, I think, more directly undercut Smith's view with respect to emotional reactions.

judgments and attitudes may thus have less to do with rationality than with *rationalization*. We sometimes demand answers from one another *as if* there were rational relations between antecedent evaluative judgments and our expressed attitudes when there aren't; instead, these "judgments" are often simply post-hoc attempts to put the sheen of rationality on what was actually a nonrational emotional appraisal (Haidt, 2001). All of this means that Smith can't necessarily get what she wants from appealing to our moral practices, as many of our actual interpersonal "answerability" demands neither seek nor receive actual answers linking attitudes to evaluative judgment. There are thus good reasons to hesitate in acceding to Smith's claims about what rational relations "we expect" to obtain between evaluations and various emotions, dispositions, and, most importantly, reasonish regard.[23]

Of course, if the dispute really comes down just to a dispute over what *we expect* the relation to be between evaluative judgment and reasonish regard, then it's an empirical claim and so is in principle empirically resolvable. I think Smith's view is really about what she thinks truly *ought* to govern our criticizable perceptual stances, independently of what we *in fact* demand of one another (as she also sometimes suggests). But if this isn't what we're committed to in our actual practices, and if there's no reason to think that evaluative judgment is the type of thing that could govern reasonish regard anyway, then her ought doesn't imply can.

7. Conclusion: Musing on Method

I have made the case for including empathic control as a type of responsibility-generating control. But this case is crucially conditional: *If* control is necessary for responsibility, then we have to include empathic control as a distinct type of control, alongside volitional and evaluative control, to account for all the things for which we hold one another responsible. But some theorists have argued that control may not be necessary for responsibility after all (see, e.g., Sher, 2006;

[23] I'm well aware that there are also empirically-driven complaints that have been made about empathy, that it is, for instance, biased and parochial (see Prinz, 2011a and 2011b; Bloom, 2013 and 2016). I have argued that these critics wildly overstate their case and that their complaints don't come anywhere near undermining a theory of moral responsibility like mine that leans heavily on the capacity for empathy and its proper exercise. See Shoemaker (2017).

see also Shoemaker, 2015: 224-25). Denying the antecedent of the hypothetical would thus obviate the need for empathic control. Alternatively, it might be thought that because empathic control doesn't have as clear a relation to agency as do volition and reasons evaluation, we should still restrict the types of control relevant to *responsibility* to the volitional and the evaluative. Perhaps empathic control is a type of control, then, but it's not a type of responsibility-delivering control.

Nevertheless, there is a theoretical advantage to taking the story I've told seriously. Doing so would respect the diversity of our responsibility practices, taking their deliverances quite seriously, while also attempting to hold onto a unifying and grounding control condition that many theorists insist upon. One could thus still say that control is required for all forms of responsibility, just allowing that there's a different type applicable to each of the three agential products (actions, attitudes, and reasonish regard). And what could unify the different types of control as *control* is that there's an analogous mechanism-based interpretation of it available for all three.

I do not aim to resolve these methodological questions, only to flag them. There is, I think, a strong prima facie case to be made that *if* we view control as necessary to responsibility, and *if* we think there is good reason to distinguish between voluntary and evaluative control, then we also have good reason to incorporate empathic control as a distinct – third – type of responsibility-implicating control. But these are big "ifs," and how we treat them will likely have radically different effects on the conclusions we come to about the nature of our responsible agency.

REFERENCES

Adams, R.M. (1985) Involuntary Sins. *The Philosophical Review* 94: 3-31.

Arneson, R.J. (2004) Luck Egalitarianism Interpreted and Defended. *Philosophical Topics* 32: 1-20.

Audi, R. (2013) *Moral Perception*. Princeton: Princeton University Press.

Baron-Cohen, S. (1997) *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.

Bloom, P. (2013) The Baby in the Well: The Case Against Empathy. *New Yorker* 20.

Bloom, P. (2016) *Against Empathy*. New York: Harper Collins.

Blum, L (1991) Moral Perception and Particularity. *Ethics* 101: 701-25.

Blum, L. (1994) *Moral Perception and Particularity*. Cambridge: Cambridge University Press.

Brink, D., and Nelkin, D. (2013) Fairness and the Architecture of Responsibility. *Oxford Studies in Agency and Responsibility* 1: 284-313.

Clarke, R. (2014) *Omissions*. Oxford: Oxford University Press.

De Martino, B., et al. (2008) Explaining enhanced logical consistency during decision making in autism. *Journal of Neuroscience* 28: 10746-10750.

D'Arms, J, and Jacobson, D. (2000) The Moralistic Fallacy: On the 'Appropriateness' of the Emotions. *Philosophy and Phenomenological Research* 61: 65-90.

D'Arms, J, (2003) The Significance of Recalcitrant Emotion (or, Anti-quasijudgmentalism) *Royal Institute of Philosophy Supplement* 52: 127-145.

D'Arms, J, (2006) Anthropocentric Constraints on Human Value. In Russ Shafer-Landau, ed., *Oxford Studies in Metaethics* 1. Oxford: Oxford University Press, pp. 99-126.

D'Arms, J, Forthcoming. *Rational Sentimentalism*. Oxford: Oxford University Press.

Darwall, S. (2006) *The Second-Personal Standpoint*. Cambridge, MA: Harvard University Press.

Dawson, G., Margaret, F.(1987) Perspective-taking ability and its relationship to the social behavior of autistic children. *Journal of Autism and Developmental Disorders* 17: 487-498.

Denham, A. E. (2017) Empathy and Moral Motivation. In Heidi L. Maibom, ed., *The Routledge Handbook of Philosophy of Empathy* (New York: Routledge), pp. 227-241.

Doris, J. M. (2015) *Talking to Our Selves*. Oxford: Oxford University Press.

Duff, R.A. (2007) *Answering for Crime: Responsibility and Liability in the Criminal Law*. Bloomsbury.

Duff, R.A. (2009) Legal and Moral Responsibility. *Philosophy Compass* 4: 978-986.

Duff, R.A. (2010) Psychopathy and Answerability. *Responsibility and Psychopathy* () pp. 199-212.

Fischer, J.,M., Ravizza, M. (1993) Introduction. In Fischer and Ravizza, eds., *Perspectives on Moral Responsibility* (Ithaca, NY: Cornell University Press), pp. 1-41.

Fischer, J.,M (1998) *Responsibility and Control*. Cambridge: Cambridge University Press.

Fischer, J.M., and Tognazzini, N. A. (2009) The Truth about Tracing. *Noûs* 43: 531-556.

Frankfurt, H. (1988) *The Importance of What We Care About*. Cambridge: Cambridge University Press.

Fritz, K. G. (2016) *Responsibility, Control, and the Nonvoluntary*. Diss. The Florida State University.

Ginet, C. (1990) *On Action*. Cambridge: Cambridge University Press.

Glover, J. (1970) *Responsibility*. New York: Humanities Press.

Glover, J. (1999) *Humanity: A Moral History of the Twentieth Century*. London: Yale University Press.

Haidt, J. (2001) The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment. *Psychological Review* 1087: 814-34.

Hieronymi, P. (2006) Controlling Attitudes. *Pacific Philosophical Quarterly* 87: 45-74.

Hieronymi, P. (2009) Two Kinds of Agency. In Lucy O'Brien, ed., *Mental Actions* (Oxford: Oxford University Press), pp. 138-62.

Hieronymi, P. (2014) Reflection and Responsibility. *Philosophy & Public Affairs* 42: 3-41.

Hobson, J. A., et al. (2009) Anticipatory concern: A study in autism. *Developmental Science* 12: 249-263.

Hobson, P., et al. (2006) Foundations for Self-Awareness: An Exploration through Autism. *Monographs of the Society for Research in Child Development* 71: 1-166.

Jacobson, D. (2011) Regret, Agency, and Error. *Oxford Studies in Agency and Responsibility* 1: 95-125.

Kane, R. (1996) *The Significance of Free Will*. Oxford: Oxford University Press.

Kennett, J. (2002) Autism, Empathy, and Moral Agency. *Philosophical Quarterly* 52: 340-57.

Klink P. C., et al. (2008) Early interactions between neuronal adaptation and voluntary control determine perceptual choices in bistable vision. *Journal of Vision* 8: 1-18.

Levy, N. (2005) The Good, the Bad, and the Blameworthy. *Journal of Ethics and Social Philosophy* 1.

Levy, N. (2014) *Consciousness and Moral Responsibility*. Oxford: Oxford University Press.

Madigan, C. (2013) *Does Thomas Scanlon's Theory of Blame Resolve the Problem Presented by Moral Luck?* Diss. Université Saint-Paul/Saint Paul University.

Maibom, H. L. (2010) Imagining Others. *Les Ateliers de L'Ethique* 5: 34-49.

McHugh, C. (2017) Attitudinal Control. *Synthese* 194: 2745-2762.

McHugh, C., Way, J. (2022) All Reasons are Fundamentally for Attitudes. *Journal of Ethics and Social Philosophy*. DOI: https://doi.org/10.26556/jesp.v21i2.1341.

McGeer, V., and Pettit, P. (2015) The Hard Problem of Responsibility. *Oxford Studies in Agency and Responsibility* 3: 160-188.

McKenna, M. (2005) Where Frankfurt and Strawson Meet. *Midwest Studies in Philosophy* 29: 163-180.

McKenna, M. (2008) Putting the Lie on the Control Condition for Moral Responsibility. *Philosophical Studies* 129: 29-37.

McKenna, M. (2012) *Conversation and Responsibility*. New York: Oxford University Press.

McKenna, M. (2013) Reasons-Responsiveness, Agents, and Mechanisms. *Oxford Studies in Agency and Responsibility* 1: 151-83.

McKenna, M. and Pereboom, D. (2016) *Free Will: A Contemporary Introduction*. New York: Routledge.

Mele, A, (2006) *Free Will and Luck*. New York: Oxford University Press.

Moran, R. (2001) *Authority and Estrangement: An Essay on Self-Knowledge* (Princeton: Princeton University Press)

Nagel, T. (1979) Moral Luck. In *Mortal Questions*. Cambridge: Cambridge University Press, pp. 24-38.

Nichols, S. (2004) *Sentimental Rules*. Oxford: Oxford University Press.

O'Connor, T. (2000) *Persons & Causes*. Oxford: Oxford University Press.

Ozonoff, S., Bruce F. Pennington, and Sally J. Rogers. (1991) Executive function deficits in high-functioning autistic individuals: relationship to theory of mind. *Journal of Child Psychology and Psychiatry* 32: 1081-1105.

Parfit, D. (1984) *Reasons and Persons*. Oxford: Oxford University Press.

Pereboom, D. (2001) *Living Without Free Will*. Cambridge: Cambridge University Press.

Pereboom, D. (2014) *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press.

Portmore, D. W. (2019) Control, Attitudes, and Accountability. *Oxford Studies in Agency and Responsibility* 6: 7-32.

Prinz, J. (2011a) Against Empathy. *The Southern Journal of Philosophy* (Spindel Supplement) 49: 214-33.

Prinz, J. (2011b) Is Empathy Necessary for Morality? In Peter Goldie & Amy Coplan, eds., *Empathy: Philosophical and Psychological Perspectives* (Oxford: Oxford University Press), pp. 211-29.

Rehfeldt, R. A., et al. (2010) Assessing relational learning deficits in perspective-taking in children with high-functioning autism spectrum disorder. *The Psychological Record* 57: 4.

Scanlon, T.M. (1998) *What We Owe to Each Other*. Cambridge, MA: The Belknap Press of Harvard University Press.

Scanlon, T.M. (2008) *Moral Dimensions*. Cambridge, MA: Belknap Press of Harvard University Press.

Scarantino, A. (2014) The Motivational Theory of Emotions. In Justin D'Arms and Daniel Jacobson, eds., *Moral Psychology & Human Agency* (Oxford: Oxford University Press), pp.156-85.

Searle, J. R. (2001) *Rationality in Action*. Cambridge, MA: MIT Press.

Sher, G. (2006) Out of Control. *Ethics* 116: 285-301.

Shoemaker, D. (2007) Moral Address, Moral Responsibility, and the Boundaries of the Moral Community. *Ethics* 118: 70-108.

Shoemaker, D. (2011) Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility. *Ethics* 121: 602-632.

Shoemaker, D. (2017) Empathy and Moral Responsibility. In Heidi L. Maibom, ed., The Routledge handbook of Philosophy of Empathy (London: Routledge), pp. 242-251.

Shoemaker, D. (2018) Cruel Jokes and Normative Competence. Social Philosophy & Policy 35: 173-195.

Shoemaker, D. (2020) Empathic Self-Control. In Alfred R. Mele, ed., Surrounding Self-Control (Oxford: Oxford University Press), pp. 384-399.

Shoemaker, D. (2015) *Responsibility from the Margins*. Oxford: Oxford University Press.

Sidgwick, H. (1981) *The Methods of Ethics*. Indianapolis: Hackett.

Smart, J.J.C. (1970) Free Will, Praise, and Blame. In Gerald Dworkin, ed., *Determinism, Free Will, and Moral Responsibility* (Englewood Cliffs: Prentice-Hall), pp. 196-213.

Smith, A. (2006) Cognitive Empathy and Emotional Empathy in Human Behavior and Evolution. *Psychological Record* 56: <http://open-siuc.lib.siu.edu/tpr/vol56/iss1/1/>.

Smith, A. M. (2004) Conflicting Attitudes, Moral Agency, and Conceptions of the Self. *Philosophical Topics* 32: 331-352.

Smith, A. M. (2005) Responsibility for Attitudes: Activity and Passivity in Mental Life. *Ethics* 115: 236-71.

Smith, A. M. (2008) Control, Responsibility, and Moral Assessment. *Philosophical Studies* 138: 367-392.

Smith, A. M. (2012) Attributability, Answerability, and Accountability: In Defense of a Unified Account. *Ethics* 122: 575-89.

Smith, A. M. (2015) Responsibility as Answerability. *Inquiry* 58: 99-126.

Strawson, P.F. (1962) Freedom and Resentment. *Proceedings of the British Academy* 48 (1963): 1-25. Reprinted in (and page numbers from) Gary Watson, ed., *Free Will* 2nd Edition (Oxford: Oxford University Press, 2003), pp. 72-93.

Taylor, R (1970) *Good and Evil*. New York: Macmillan.

Todd, P. (2016) Strawson, Moral Responsibility, and the 'Order of Explanation': An Intervention. *Ethics* 127: 208-240.

Toplak, M. E., et al. (2010) Decision-making and cognitive abilities: A review of associations between Iowa Gambling Task performance, executive functions, and intelligence. *Clinical Psychology Review* 30: 562-581.

Vargas, M. (2005) The Trouble with Tracing. *Midwest Studies in Philosophy* 29: 269-291.

Vargas, M. (2013) *Building Better Beings*. Oxford: Oxford University Press.

Wallace, R. J. (1994) *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press.

Watson, G. (2004) *Agency and Answerability*. Oxford: Oxford University Press.

Watson, G.(2014) Peter Strawson on Responsibility and Sociality. *Oxford Studies in Agency and Responsibility* 2: 15-32.

Webster, A, (2021) Socially Embedded Agency: Lessons from Marginalized Identities. *Oxford Studies in Agency and Responsibility* 7: 104-29.

Williams, B. (1993) *Shame and Necessity*. Berkeley, CA: University of California Press.

Wolf, S.. (1987) Sanity and the Metaphysics of Responsibility. In Ferdinand Schoeman, ed., *Responsibility, Character, and the Emotions* (Cambridge: Cambridge University Press), pp. 46-62.

Zhang, Xue, et al. (2017) The interaction of perceptual biases in bistable perception. *Scientific Reports* 7. Doi: 10.1038/srep42018.