

## **The Social Constitution of Agency and Responsibility: Oppression, Politics, and Moral Ecology**

Manuel R. Vargas

**ABSTRACT:** Culpability under oppression is puzzling. On the one hand, it seems callous to insist that someone's being subject to oppression is never relevant to their culpability. On the other hand, responsible agency can and does exist under systematic disadvantage, and insisting that oppression always undermines a person's culpability is both too forgiving and disrespectful to those agents. One philosophical challenge that grows out of reflecting on these matters is whether there is a way to reconcile the thought that our agency is socially constituted, or a product of circumstance, with the thought that some psychological configurations rightly underpin our condemnatory practices while others do not. This paper offers a framework for thinking about these issues, and in particular, the question of culpable action under conditions of oppression. The account maintains that oppression matters for culpability in some cases and not in others, in part because culpability is a function of both the properties of agents but also the social (and political) context that shapes our agency. One consequence of this approach is that the moral ecology of our agency—the circumstances that support and enable morally valuable forms of agency—matter a great deal for agency and responsibility.

When people are subject to oppression, does that fact undermine their culpability for wrongdoing?<sup>1</sup> No uncomplicated answer is appealing. On the one hand, it seems callous to insist that someone's being subject to oppression is never relevant to her culpability. On the other hand, it seems implausible and disrespectful to insist that no one is ever responsible for wrongdoing under oppression.

One skeptical view is that this puzzle about oppression is the product of a deeper problem. The conviction that people can be responsible under oppression is a recalcitrant conviction grounded in the naïve belief that people are free in a way that allows them to transcend their circumstances. Similarly, the exculpatory impulse that emerges when we think about oppression is a species of the more general thought that *all* actions are conditioned in the same sort of way, as products of prior causal forces. On this view, the puzzle about culpability under oppression is just the conflict between our naïveté and our recognition that we are products of circumstance.

If this skeptical worry is familiar, so is the reply. Optimists insist that there is a way to ground the difference between culpable and non-culpable action. They typically appeal to some privileged bit of psychological functioning: the possibility of acting from values, commitments, or rationality; the reasonability of asking people to justify themselves; or the putatively theory-neutral commitments of affective responses to wrongdoing. In the context of thinking about culpability under oppression, however, there is something *prima facie* dissatisfying about the optimist's strategy. To point to psychological features as the basis of responsibility misses the force of the worry that oppression undermines culpability. The skeptical worry *just is* that our psychologies are shaped by social context in a way that makes us not culpable.

There are two ways forward. One is to insist on the philosophically fraught possibility of our being able to transcend our circumstances in some robust way. A second, perhaps less intuitively appealing approach, is to show how to acknowledge that our agency is socially constituted without surrendering the distinction between culpable and non-culpable actions. This

---

<sup>1</sup> *Oppression* is an elastic term in the wider philosophical literature, admitting of wide differences in its characterization (Haslanger et al. 2015). Let *oppression* be the property of unjust or immoral treatment, social relations, or distributions of opportunities, when it is produced by immoral or unjust social and political arrangements (cf. Cudd 2006, 20). The basic idea can be refined in distinct ways, with varied emphases on the role of institutions (Cudd 2006: 50), subordinate group membership (Haslanger 2012), or oppression's characteristic psychological, social, or material conditions (Young 2004). What follows is compatible with standard refinements of the idea of oppression.

chapter pursues the second path.

An idea that animates some of what follows is that we tend to think about culpability in the wrong way, as almost exclusively settled by intrinsic features of the agent and her actions. Culpability is a function not just of agents, but of the agent's relationship to circumstances and the governing normative practices. The foundations of culpability are as much political and normative as metaphysical. Oppression is particularly interesting because it has varied effects on culpability: sometimes it eliminates it, sometimes it makes it harder to do the right thing, and other times, it enables responsible agency. The upshot is that the nature of culpability under oppression—and what we can do about it—is complex.

Section one outlines the basic puzzle and reasons for rejecting the idea that we possess a transcendent form of agency that obviates the need to explain how social forces might affect culpability. Section two provides an account of the basis of moral culpability that accommodates the fact of the social constitution of our agency while specifying the normative function of responsibility practices. On this picture, responsible agency is socially constituted in two ways. First, the acquisition of our values and dispositions, and how they are maintained is a fundamentally social matter, continuously shaped by social feedback. Second, facts about what social practices are effective at fostering moral considerations—facts outside the head of the agent—fix what capacities matter for culpability, and determine whether a given agent's psychology can be rightly said to be able to respond to moral considerations. Section three applies this account to culpability under oppression, arguing that oppression can affect our ability to respond to moral considerations as well as the standards that determine what it is for us to be able to respond to moral considerations. The final section explores the implications of this picture for larger questions about the social and political ecology of culpable agency. Once we see culpability as partly a function of the social constitution of agency, we have reasons to reflect on the limits of social and political interventions that structure that agency. In sum, agents can be culpable under oppression, and this is something we have some control over, but the particulars matter.

### **1. Agency in Circumstances**

I began by noting that it can seem both disrespectful and too forgiving of oppressed people to maintain that they are not culpable for wrongdoing under oppression. The skeptic will object

that, respectful or not, it is false that people can be responsible under oppression. I reject the skeptical view, but it is instructive to begin by considering an idea that animates it, namely, the evident tension between two families of truisms about the agency required for culpability.

One set of truisms about individual agency includes the following: People are ordinarily agents; sometimes they act intentionally, other times they don't; some actions are freely undertaken, others are not; some actions deserve moral praise and blame, and others do not. A second set of truisms concern the relationship of context and agency. This set includes thoughts of this sort: what people do is a matter of their circumstances; circumstances matter both for what we care about and what we can do; what seems valuable or appealing to an agent is a product of upbringing and socialization; what powers a person has in a given situation are products of history and forces outside of the control of the agent.

It is unclear how both sets of ideas could be simultaneously true. Our commonsense ideas of culpable agency—that agents are at fault, that they deserve blame and condemnation for wrongdoing—seem to require some independence from circumstances. When the effects of circumstances are particularly salient, as when one is stressed or subject to some unusual pressure, we tend to excuse the effects of those circumstances. That exculpatory impulse can seem like the thin edge of a wedge. Once we recognize the fact that who we are and what we do is a product of either antecedent forces or circumstances over which we have no control (as we do when we forgive someone for being stressed or exhausted), it is unclear how this is different in kind from any circumstantial source of our dispositions and behavior. Accountability threatens to dissipate in the face of appeals to circumstance.

The challenge here is not always one of clear contradiction. Rather, it seems to be a matter of one explanation trumping or undermining another (cf. Björnsson and Persson 2013). In our interpersonal lives, we tend to act as though our ideas about individual agency are fundamental and secure. When we shift our attention to context, though, the fact that agency seems to be a product of social (and other) kinds of circumstances somehow makes the truisms about individual agency seem less evident. In the social sciences, this tension is often characterized in terms of the structure-agency problem. Especially (but not exclusively) in sociology, a persistent challenge is how to square explanations that emphasize the explanatory or causal role of social structures with the idea that individual agents make free choices. In philosophy, variants of the problem tend to emerge in debates about moral luck, free will, moral

responsibility, and autonomy. Determinism, mechanism, socialization, or the social structuring of an agent's option set is taken to threaten the possibility of various ideals of agency, including morally responsible agency (e.g., Pereboom 2001; Greene and Cohen 2004).

Some caution is in order. We sometimes talk of agency, full stop, as though it were a single thing, the contours of which we all agree upon. This way of talking can mislead, both about the variety of forms of agency and about the particular challenges that threaten distinct forms of agency. There are convictions about agency that are unaffected by attention to circumstances. Creatures resolve practical challenges with different kinds of deliberation: the squirrel decides whether to go for the acorn, the writer decides whether to keep typing, and the reader decides whether to take notes. These sorts of facts—facts about some pedestrian forms of agency—aren't undermined by our acknowledging that these agents are part of a larger social or environmental context. The agency of the squirrel-as-acorn-pursuer, person-as-typist, and reader-as-note-taker is entirely compatible with thinking each is structured by, and a part of, external forces. Contrary to one reading of Nagel, agency does not always shrink “to an extensionless point” when we consider context and antecedents (1979, 35).

Redirecting our attention to wider social or environmental forces can diminish the salience of particular psychological facts about given agents. Diminished salience is only that, however. It is not the more radical and unwarranted claim that there is no genuine agency—that agents do not believe, desire, decide, intend, and attempt. What is more plausibly threatened by attention to circumstances are those forms of agency that make a claim about normative standing. Whether someone is free in a politically or morally relevant sense, whether we should take seriously someone's preferences in formulating our plans, whether we can appropriately praise and blame, whether someone acts autonomously—we can become uncertain about such questions when we highlight the role of context. Where values, sensibility, and the perception of reasons are structured or constituted by culture, norms, and values, and where the range of available acts is constrained by institutions, resources, and wider social meanings, it is harder to be confident about our normative assessments of agents.

Even if we were convinced that, in general, agency can be suitably distinguished from the wider causal fabric in a way that might support responsibility attributions, the internal structure of our practices show some concern for the social origins of our motivations, values, and judgments. At the first pass we tend to suppose that an adult of ordinary knowledge and

capabilities is culpable for knowingly putting a child at risk of unnecessary injury. So, if a soccer parent harms a child by insisting on unreasonable athletic performance—e.g., insisting on the child playing through a significant injury—we are inclined to hold the parent responsible. (We are also often prepared to find fault for modestly indirect effects of parenting choices on the maturing child.) However, that initial presumption of culpability is vulnerable to exculpatory pressures *internal to the practice*. Suppose we learn that our soccer-crazed adult grew up in a sports-obsessed environment, one that downplayed the seriousness of the risk to the child while glorifying the risk-taking of playing while injured. For many of us, this additional bit of information makes the initial certainty that the parent is fully culpable at least somewhat less certain. To the extent that one acquires normative standards by cultural osmosis, and one is in a cultural milieu that discourages and even attempts to disable critical reflection on those standards, it is less obvious that the offending adult is as culpable as our initial reactions tend to suggest. The fact of the always-lurking exculpatory element of context invites further explanation.

There is a more general phenomenon at work here. When our values, knowledge, or moral convictions are at some remove from those we are considering condemning, there is some tendency—for at least us, here and now—to be unsure how to proceed in our assessments of blame. We might still find their behavior repugnant, ill-considered, or reprehensible. The strength of these reactions and what follows presumably reflects, among other things, our sense of the degree to which the offender is a member of our in-groups, the nature of the perceived offense, and so on. However, our confidence that some action is morally reprehensible can rest (perhaps with some discomfort) alongside our uncertainty about the culpability of those performing it (Watson 1987). Wrongfulness and culpability come apart.

A good deal of our moral sensibility is socially constituted. However, is not merely that social influences sometimes bypass our rational, valuing agency but that these influences *just are* the things that generate and sustain our motivations, cares, values, and habits of regard in the first place. That's why the typical approach of optimists about responsibility—rooting distinctions of culpability in privileged bits of psychology, whether in desires, values, rational capacities, commitments, or judgments—seems unresponsive to the thread of thinking that seems to animate one part of our attitudes about oppression. Once we attend to it, the usual way we

acquire values and commitments seems to undermine culpability.<sup>2</sup>

Here some will interject with the claim that it is the possession of free will that allows us to disregard questions of social context.<sup>3</sup> If an appeal to free will is supposed to help explain why agents can be influenced by context but still culpable, then free will must be the sort of thing that licenses our praising and blaming of one another—even in the face of the evident fact that our choices, values, and aspirations have been subject to social and other influences.<sup>4</sup> Transcendent agency is roughly the idea that agency is not importantly constrained by or limited to features of circumstance or experience. The idea that we have a kind of “contra-causal” freedom, perhaps coupled with something like substance dualism, is one way of giving expression to this elusive thought.

The appeal of this picture of agency is clear. If we do have such freedom, then the social dimensions of the structure and exercise of our agency is mostly irrelevant, precisely because it can be transcended or overcome. Moreover, the view seems to have deep roots in aspects of ordinary convictions about agency (Vargas 2013a; Nichols 2015).

Alas, the idea of transcendent agency is more aspirationally alluring than philosophically satisfying. First, it is not clear what basis there is—or could be—for transcending anything. On standard philosophical and scientific models of agency, anything we do is a product of circumstance, biology, physical structure, or the interaction of these things. To the extent to which these forces are not in our control, it is not clear how anything we do could count as transcending these influences, absent some special story that is at odds with most current

---

<sup>2</sup> This account is neutral on whether positively valenced notions—moral credit and praise—can be modeled on assessments of (negative) culpability.

<sup>3</sup> Let *free will* pick out a power or capacity distinctive of creatures properly subject to evaluation in terms of praiseworthiness and blameworthiness. Nothing in this way of framing matters precludes the possibility of alternative understandings of the term. Ongoing disagreements about what free will *is*, what it *means*, and what it *refers to* are partly artifacts of the wide range of philosophical concerns that have been pursued under the guise of theories of free will. Accounts of distinct aspects of agency—deliberation, the phenomenology of freedom, causal origination, authorship, the ability to do otherwise, and so on—each plausibly fix on some property that might reasonably be characterized as free will. It is not obvious that they all come to the same thing (cf. Vargas 2011).

<sup>4</sup> These issues are plausibly related to both *the socialization problem* in the autonomy literature, especially as it is concerned with relational notions of autonomy (cf. Mackenzie and Stoljar 2000; Westlund 2009; Oshana 2014, 2015) and to worries about adaptive preferences (Stoljar 2014). For present purposes, I’m neutral on the relationship of moral culpability to various notions of autonomy.

scientific and philosophical views. Second, it is not clear what the basis for decision-making would be, even if we could transcend the particulars of our context, biology, or underlying physical structure. It is not obvious why decision-making entirely disconnected from the discernible features of the context of decision would be an appealing picture for the basis of culpability. Transcendent agency would also be opaque to us—inasmuch as it would be radically disconnected from our understanding of the physical, causal world, it is not clear how we could understand its basis, when it occurs, or whether someone is systematically incapable of such transcendence for some or another reason.<sup>5</sup>

It is not irrelevant that the various brain, cognitive, and behavioral sciences suggest that our everyday notion of agency is systematically mistaken in a number of different ways, and that may include our sense of our being transcendent agents. Conscious, reflective agency seems to play a considerably lesser role in what we do than ordinary conceptions of agency presume (Nahmias 2010; Levy 2014; Doris 2015). We are prone to invisible biases whose operations are not always readily discernible, and seemingly irrelevant features of situations and contexts propel and sometimes disable agency in ways agents fail to recognize and would tend to disavow if informed (Doris 2002; Nelkin 2005; Nahmias 2007; Vargas 2013b; Doris 2015; McKenna and Warmke forthcoming). This is not to make the radical point that conscious control is illusory, that we never understand our motives, or that self-aware agency is impossible to attain. There are various reasons to think that the most radical versions of these views are implausible (Mele 2009). It nevertheless remains true that we oftentimes have an uneasy grip on why we do what we do.

What I take to be the general lesson of contemporary cognitive science is this: we underestimate the degree to which our actions are structured by situational pressures and automatic processing that operates below the level of conscious awareness. In turn, this suggests a kind of challenge to views that insist on our transcendence—even if we really did sometimes transcend our circumstances, biology, and physical relations, it is clear that automatic processing, regular, discernible patterns of bias, situational effects, and so on are the order of the day. Given

---

<sup>5</sup> One motivation for some forms of libertarianism about free will is presumably the desire to capture the idea of transcendence. It is unclear the extent to which all or any philosophically credible libertarian view adequately captures a potentially more basic and radical appeal to transcendence. Libertarianism remains a controversial view, and I have argued against it elsewhere (e.g., Vargas 2004, 2013a, 2016).



that most exercises of agency don't suggest a radical independence from the causal and social order, we are back to the basic challenge: to the extent that we are a part of—embedded in, we might say—a social order, and given the forms of agency that are compatible with a causal, naturalistic conception of humans, can we show that people are culpable?

The issues here are complex. Does social deprivation (in the form of poverty or social stigma, for example) impair responsible agency? Is otherwise culpable action subject to mitigation when performed under conditions of oppression (as when a battered spouse kills the batterer)? Does unjust or repeated exposure to risks or harms change the moral significance of some otherwise culpable acts?<sup>6</sup> The philosophical literature on freedom and responsibility has been mostly silent on these questions, but if we accept that culpability has a relationship to the social constitution of our agency, these questions are not peripheral.

It would be a mistake to suppose that political and legal considerations exhaust the way in which our agency is socially structured. There is a web of more broadly social categories and roles that structure which actions seem appealing and how we think about ourselves (Hacking 1995; Haslanger 2012), which values are easier to live out, and the control one has over the perception of one's intentions (Bierria 2014). The wrong sorts of socialization seem to undermine one's suitability as a target of the reactive attitudes (Wolf 1987; Watson 1987). If one goes on to think that, in general, our agency is socially scaffolded (Vargas 2013a; McGeer 2015; Doris 2015; Huebner 2016), or dependent on interaction and feedback from other agents, then both the presence but also *the absence* of certain kinds of social feedback will affect whether someone is a responsible agent. The issue here is not just whether one is responsible or not. Sometimes the stakes are subtler. A particularly insidious way for the social world to be structured is where, perhaps invisibly, the context makes it particularly difficult to exercise one's responsible agency.

One promising idea is that it is better to think of sociality as an enabling condition of our agency, rather than a constraint. Vindicating this idea is no easy task. Beyond accommodating the idea that who we are and what we value are frequently given by, or at least structured by, social context, an adequate account will also need to accommodate the idea that what *opportunities*

---

<sup>6</sup> Questions about degree of blame and the moral valence of individual acts might seem separable from questions about whether someone is a responsible agent, i.e., an apt target of moral responsibility. Although questions of responsible agency and degree of culpability (and indeed, the relevant moral valence of the act) are conceptually independent, I will argue that the way social facts can undercut responsible agency turns out to matter for these other questions, too.

we have are often socially structured. Collectively, these thoughts suggest that we do poorly to approach responsible agency in atomistic, or individualist terms. If the nature of our agency and the social structure of our opportunities are products of, among other things, political institutions and practices, then responsible agency is partly a matter of how our social, political, and even ideological circumstances are arranged. Call this *the challenge of social circumstances for moral responsibility*.

Let's turn to how we might answer that challenge.

## 2. Agency Cultivation

Here, I propose a way of balancing our commitments to the idea of responsible agency with facts about the social structuring, scaffolding, and constitution of that agency. There are two desiderata governing what follows. First, the ambition is to offer a metaphysically modest account of moral responsibility that allows for *the social constitution of agency*. Second, the account must provide an *adequate normative basis* for culpability judgments. I begin with the first desideratum, concerning the social constitution of agency.

Non-derivative culpability for wrongdoing is a function of at least three things: *rational capacities, psychology, and environment*. Culpable agency involves a suitable ability to recognize and respond to what moral considerations there are.<sup>7</sup> However, our psychology involves bundles of relatively circumscribed cognitive and affective processes (some more abstract and deliberative, some more concrete and automatic). So, following standard views in the cognitive sciences, I will maintain that our psychological dispositions are less cross-situationally stable, general, or content neutral than commonsense views tend to suppose. Instead, they are clusters of dispositions that function best in relatively circumscribed contexts. As a result, the boundedness of our rationality means that our rational capacities are better- and worse-suited to specific and circumscribed environments. In short, our rational agency—and thus, our culpability—is a product of the fit between our psychologies and our environments.

These three foundations of culpability (our rational agency, its psychological underpinnings, and the environment) have a dynamic relationship to one another. We can “train up” cognition for particular environments. Environments can foster particular patterns of cares

---

<sup>7</sup> For present purposes, I am neutral on how reasons are related to our cares and commitments.

and commitments that shape what agents perceive as reasons. We can also restructure our environments to better exploit our cognitive and affective dispositions, and to better express and realize our cares and commitments. As Hurley has noted, both higher-level cognition—e.g., explicit meta-representation—and lower level, more automatic processes have a “public ecology” that fosters and sustains situated rationality (2013, 206). The socially scaffolded nature of our agency is thus not an incidental but central feature of the kinds of creatures we are.<sup>8</sup>

Here, though, the animating challenge of social circumstances becomes central. If the social constitution of our agency is what makes oppression look exculpatory, then how does a socially-situated picture of rational moral agency solve the problem? To answer this question, we must turn to the second aspiration for the present account. That is, we need to show how the proposal is *normatively adequate*, or capable of doing the normative work we need for a theory of responsibility. The account succeeds if it can show how socially constituted, circumstantially variable features of agency can license, justify, or make normatively apt our culpability judgments and their characteristic attitudes.

In what follows, I offer an account with three important elements: (1) a justification for responsibility attitudes and practices (the “responsibility system”); (2) an account of the form of agency required for culpability; and (3) an explanation of how this is compatible with an ecological, non-transcendent picture of agency. Together, these elements provide the resources for showing that culpability is compatible with the social constitution of agency. I’ll say more about each element in order, but an initial sketch of the overall picture may be useful.

On this proposal, the underlying function of our responsibility attitudes and practices is in the cultivation of a certain form of agency, one that recognizes and responds to moral considerations. Call this justification of our general system of responsibility practices the *agency cultivation model*. However, the norm-structured practices of moralized praise and blame only rightly apply to agents that are suitably capable of responding to moral considerations. Importantly, the involved capacities of responsiveness are *circumstantialist*, or sensitive to a distinctive set of social and normative considerations that vary across circumstances. This

---

<sup>8</sup> There are more and less radical ways to accommodate the idea of the circumstance-specific powers of our agency. As will become clear, I favor an approach according to which the normatively significant powers of agency alter by context even if the material substrate, as it were, is stable. However, the basic theoretical demand is more modest than that. What is required is that a theory of responsible agency has *some* way of capturing the way circumstances seem to impinge on agency.

conception of capacity is tied to the justifying aim of the practice: the relevant notion of capacity is the one that, were it adopted and internalized by the relevant agents, would be at least co-optimal for improvements in the agency of those governed by the practice. Together, these ideas can provide a unified story about whether and when agents are culpable for wrongful action under conditions of oppression.

Let's begin with the justification of the responsibility attitudes and practices. Social practices that admit of a distinction between the justification of the practice and the content of the particular rules of the practice provide a helpful model for thinking about the relationship of culpability to the justification of responsibility practices. For example, consider a system of foul calls in a particular sport. A standard justification for having foul calls in a sport is that it helps protect the players, and, perhaps not unrelatedly, the enjoyment of the fans. Internal to a game, though, fouls are assessed on the basis of the foul rules. Whether a player deserves a foul is not determined by the referee in light of whether it is appealing to the fans, or whether the calling of the foul will in fact protect a given player. Rather, the basis of the foul is the rules that govern the game and the players in it. When one violates those rules in particular ways, one deserves a foul. To be sure, a given referee's decision about whether to call the foul can be subject to a range of epistemic and pragmatic considerations. Sometimes it is unclear whether a foul was committed. A minor but genuine foul in the waning moments of a game that is already decided may not seem worth the call. Even so, the propriety of the foul, considered as a foul, is settled by the actions of the players and the applicable rules. Importantly, it is *not* a matter of whether making the call in a particular instance contributes to the teleology of a system of foul calls.

Culpability—or blameworthiness—functions in a way analogous to a system of foul calls. That is, an agent is blameworthy when she has failed to act with due moral concern and is the right sort of agent for the application of those norms. Like a system of foul calls in a sport, the overall practice of moral blame has a teleology to it. This teleology is not operative at the level of first-order determinations of culpability (or a foul). In both practices, considerations about the justification of the system of norms and practices may be entirely invisible to participants in the practice. However, what makes a system of blame justifiable (when it is) is that it cultivates moral considerations-sensitive agency. Call this the *agency cultivation model* of moral responsibility. To the extent to which a system of moralized praise and blame helps refine and extend our ability to recognize and suitably respond to moral considerations, it has a justification. To the extent to

which it fails to support this form of agency—and absent some further justification—it fails to be normatively adequate.<sup>9</sup>

The norms of blameworthiness apply to some agents and not others. Infancy and some varieties of mental impairment are commonly regarded as conditions that preclude membership in culpability practices. On this account, some baseline ability to recognize and suitably respond to moral considerations is a prerequisite for the applicability of the blameworthiness norms. When one has that baseline ability, then one is a responsible agent, i.e., a proper subject of the norms of responsibility, and a candidate for culpability. When one fails to have the requisite ability—e.g., because one cannot recognize the relevant moral reasons, or because of some systemic volitional incapacity to that class of considerations—then one is not a responsible agent, and thus, incapable of culpability.<sup>10</sup>

Crossing that threshold of ability is part of what makes someone properly (or if you like, truly) responsible, as opposed to it being merely useful to blame that person. With children, we oftentimes feign blame as a tool for teaching children about what is blameworthy. Ersatz blame is one of the ways we nurture and develop the capacities required for genuine blameworthiness, and thus, help foster genuine blameworthiness over time. However, blame (of the sort that expresses a judgment of culpability) is distinguished from moral education in part by the blamer believing the blamed to be an agent of the right sort. Where the blamed agent is missing the requisite set of recognitional and volitional capacities, the agent is not culpable. So, any blaming (or feigning of blame) requires some other basis.

A consequence of this approach is that culpability is “patchy.” One might be a responsible agent with respect to some considerations and not others. The same person might be responsible with respect to those considerations only in some circumstances and not others. Responsible agency is therefore not an always on or always off matter, and one’s suitability for blame can shift in highly context-sensitive ways. Borderline cases or atypical psychologies raise a range of puzzles for ordinary practices. However, where there is a suitable degree of sensitivity to

---

<sup>9</sup> See Vargas 2008 and Vargas (2013a) for more detailed statements of this approach. Related accounts include Arneson 2003 and McGeer 2014, 2015.

<sup>10</sup> Here, and throughout, I am bracketing cases of derivative culpability, i.e., cases where one’s present responsibility isn’t grounded in occurrent features of the agent but instead derives from a prior instance of responsibility.

the relevant moral considerations, one is a responsible agent, thereby potentially culpable for wrongdoing.

On this picture, our rationality is bounded or tied to circumstances, and more generally, agency is structured by contextually or “ecologically-situated” dispositions to respond to information in the environment. The difficult issue is how to understand claims that agents *can* respond in the relevant ways in the identified contexts. What sort of threshold of sensitivity to moral considerations would constitute the possession of the capacities required for culpability? Answering this question has been a long-standing difficulty for accounts of free will and responsibility that are intended to be compatible with the thesis of determinism. An important part of the appeal of transcendent pictures of agency is that they seem to provide more principled answers to this question than compatibilist proposals.

We can deliver an account of (non-transcendent) capacities that fits with this ecological conception of agency. Let’s assume that culpability presumes agency that is sensitive to moral considerations. So, the issue is how to specify that sensitivity or capacity in a way that can be integrated into the agency cultivation model. Here’s the solution: for a given moral consideration, the relevant degree of sensitivity is determined by whatever conception of capacity that an ideal observer would pick as best for helping imperfect creatures like us to do the right thing, over time, and in the actual world. Call that conception of capacity the *responsibility-relevant notion of ability*.<sup>11</sup>

The responsibility-relevant notion of ability can’t be so strict as to entail that no one has any possibilities other than what they do (as, for example, a conception of ability on which people can only do one thing under determinism). That wouldn’t enable a well-functioning system of blame to achieve its effects. Similarly, the notion of capacity can’t be so general as to entail that everyone always is blameworthy for everything causally connected to their choices. Such a conception of ability would make it impossible for agents to reasonably manage the risks of

---

<sup>11</sup> More precisely: the responsibility-relevant capacity is the one that, were we to internalize that conception of ability, it would generate practices of moralized praising and blaming in the actual world that are at least co-optimal for agents recognizing and suitably responding to moral considerations in the likely circumstances of action, given our psychologies and the existing socio-historical circumstances. This capacity allows us to specify a modal profile for agents, which we can construe in terms of possible worlds. Roughly, the observer-specified notion of capacity tells us what kinds of worlds count as relevant. We look to those worlds, and if in a suitable proportion of those worlds the agent responds in the relevant way to the relevant considerations, we say the agent has the relevant capacity. For details, see Vargas (2013, 213-228).

culpable choices. So, the relevant notion of capacity has to be a Goldilocks notion of capacity: not too strict and not too lax, or, just right for enabling a well-functioning system of blame.

One can get a sense of how this framework can help us understand the way oppression affects culpability by considering the idea of degrees of culpability. On this account, the ability to recognize and respond to moral considerations is a scalar matter. However, there are plausibly three distinct parts of that continuum that matter for culpability. First, there are agents who fall below the requisite degree of sensitivity to moral considerations in a circumstance and are therefore not culpable. Second, there are agents who cross the requisite threshold so that they are properly subject to the norms of blame, and thus, are culpable for the relevant bit of wrongdoing in the considered context. Among these agents, there is a further division codified in common sense, between agents whose wrongdoing is mitigated because of morally significant difficulty in exercising the relevant capacity, and those who have no such mitigating element.

The present approach has the tools to capture this idea. The ability of an agent to respond to a relevant moral consideration is a function of how the observer's conception of ability maps the agent's intrinsic properties on to a modal profile. This modal profile determines whether the agent has the ability to recognize and respond to the relevant considerations, and it comes in degrees. Agents whose modal profile just crosses the required threshold for culpability are distinct from those whose profiles cross the requisite threshold by a considerable margin. This difference suggests a useful distinction between the second and a third level of ability to recognize and respond to moral considerations. The latter, third level of responsiveness concerns modal profiles where it can be said that there is no morally salient difficulty in complying with the relevant moral norm. Variation in the quanta of blame directed at agents within this third level is, presumably, primarily a matter of the significance of the operative moral consideration and the quality of the will evinced by the agent in the considered context.<sup>12</sup>

The metaphysics of ability talk is not everyone's cup of tea. Nevertheless, it provides something too many compatibilist accounts lack: a specification of what it means to say someone had the ability to respond to moral considerations in a context. Modeling the metaphysics of

---

<sup>12</sup> Moral considerations may not all be on a par with respect to the modal profile they have. In the case of especially important moral considerations, the effects of blame may be particularly important in orienting our attention and shaping our moral concerns. So, the threshold of sensitivity for being a responsible agent may be lower for such considerations. In contrast, for less important moral considerations, the threshold may be higher simply as a matter of managing reasonable demands on our agency.

culpability-relevant ability in a way tied to a concern for achieving and sustaining moral considerations-sensitive agency blocks the worry that compatibilists are helping themselves to a cheat (i.e., a story about a non-libertarian conception of ability or capacity). It also provides an explanation of why capacities understood in this way matter for the appropriateness of blame, and this will matter for how we understand the significance of oppression for culpability.

I now turn to how this picture connects with the social constitution of our agency. Let's begin with the idea that we value being regarded as competent at navigating the social world in light of moral considerations. To be seen as incompetent in this way is to lose an important kind of social standing, and considerations like these plausibly underpin our willingness to be called to account (cf. Raz 2011). Given that one of the best ways to be reliably seen as morally competent is to be *actually* competent, practices that foster and extend such competence are important. The present account captures both our interest in being a certain kind of agent, and why culpability would be sensitive to whether, in a given situation, we meet some standard or expectation of competence at recognizing and responding to operative moral considerations. We also have a robust interest in *other* agents being sensitive to moral considerations. Our engagements with other agents become incredibly fraught when we think they are insensitive to moral considerations (or, at least, what we take to be moral considerations.) Individually and collectively, we have a pervasive interest in the kind of agency at the heart of the agency cultivation model.

Notice that all of this is compatible with, and even depends on, the socially constituted nature of agents. Blameworthiness does not require a transcendent self, detached from the causal order. Culpability for wrongdoing depends on agents, the narrow circumstances of the action, and the justified norms, which in turn shape the practice-dependent specification of the abilities that matter for culpability. Changes in each of these elements can alter culpability. Holding fixed the circumstances and wider situation, a change in the agent's intentions or the psychological structures of agency can alter culpability. Holding fixed the intrinsic properties of an agent, the agent's culpability can change if the circumstances of action are altered. Coercive circumstances, for example, can alter the culpability of an agent even when the agent's intrinsic features remain the same. Finally, even if we hold fixed the intrinsic properties of the agent and the narrow circumstances (for example, circumstances of moderate coercion), wider social or cultural contexts can ground different norms concerning whether caving in to that coercion is culpable.



The ideal observer's judgments about those norms will vary depending on what particular norms (and what specifications of capacity) best achieve the justifying aim of the responsibility system. Thus, the account offers a principled basis for distinguishing between culpability in some circumstances and not others, without denying the possibility of constitutive luck and luck in circumstances.

Culpability is not a function of whether some bit of blaming will in fact produce the justifying good in that context. The underlying normative structure of this account is teleological, but not in the way of widely-repudiated consequentialist accounts of yore. That is, it does not presume that individual judgments of responsibility are exclusively forward-looking, or only applicable to the extent to which they modify the downstream attitudes or behavior of a given agent. Akin to a system of foul calls, there is a forward-looking basis to have systems of responsibility practices with backward-looking standards of evaluation. Plausibly, those norms of blameworthiness are concerned with quality of will, or failures of to meet expected standards of moral concern. Culpability is thus a function of an agent's action, plus (typically backward-looking, quality of will-focused) norms of blameworthiness. And all of this presumes that agents have the responsibility-relevant sense of ability, specified in a way that enables a practicable and justified system of blame.

As in the case of the relevant notion of capacity, the standards of culpability are somewhat coarse-grained, on pain of making the norms and practice of responsibility too difficult to internalize and deploy. A system with a more direct and fine-grained tie to a particular justifying outcome—for example, the building out of that agent's capacity by blaming in that particular instance (e.g., McGeer 2014, 2015)—imposes considerable cognitive demands on agents. Moreover, it introduces greater unpredictability into our social practices in a way that undermines the prospective control an agent has over whether he or she is blamed (Schmidt in progress). The advantage of a view that distinguishes the justification of responsibility practices from the particular norms that achieve the specified end is that it avoids the omnipresent tyranny of instrumental blaming.

The teleological aspect of this account, even in the form at work here, can still give pause to those who feel the intuitive pressure to think that moral responsibility must be the kind of thing that is grounded in non-contingent features of the world, or even rooted in a radical kind of human freedom that sets humans apart from the rest of the causal order. Despite the pull that

transcendent pictures of agency may have on many of us, such accounts are not necessary to justify and make sense of familiar practices of moralized praising and blaming. The present account is a proposal for how we should go forward given what we now know about our agency, and it shows how practices like ours can be in good stead. We need not substitute “moral sadness” for full-throated culpability-imputing blame, as some views would have it (cf. Pereboom 2001). A sufficiently nuanced teleological account has the resources to capture the everyday phenomena of actual, messy, human praising and blaming that sent some philosophers searching for a picture of transcendent agency and left others seeking a diluted or even skeptical picture of responsibility.

### **3. Responsibility and Situated Agency**

Culpability for wrongful acts arises when responsible agents (i.e., agents with the requisite capacities to recognize and respond to moral considerations) fail to exercise moral considerations-responsive capacities in the morally appropriate way. Estimates of blameworthiness for wrongdoing can be also construed as evaluations of an agent’s quality of will. Where a responsible agent fails to conduct herself in a way that reflects due concern for morality, there is blameworthiness. For agents with the requisite capacities, the agent’s quality of will is partly a function of the quality of the options and the norms or expectations are properly applicable to those kinds of situations.<sup>13</sup> Where the options in those circumstances provide no morally adequate actional possibility to the agent, and the absence of those options are no fault of the agent, then the agent is not culpable. With the foregoing account in hand, we can now address the relevance of oppression to culpability and the distinctive way the appropriateness of blaming can be asymmetric.

Oppression can render invisible the wrongfulness of an act to both oppressor and oppressed. It can also obscure the moral significance of acts in oppressive situations. Thus, oppression seems relevant to the culpability of the oppressed, but also to the oppressor (Calhoun 1989). Given an account of culpability that requires that agents have the ability to recognize the

---

<sup>13</sup> Recall: the applicable norms will be those relatively general norms that tend to cultivate the right kind of agency, holding fixed various psychological and social features of the community to whom the norms apply. The question of how to individuate options is complex, although the relevant conception of options allows for the possibility that two acts may be behaviorally similar but deliberately distinct options, which in turn affects evaluations of the agent’s quality of will. This approach is functionally similar to the “fair opportunity” approach developed by Brink and Nelkin (2013).

relevant moral considerations, the invisibility of the wrongfulness of an act, because of oppression, can seem to exculpate both the oppressor and oppressed. This result is worrisome in each case. As noted at the outset it seems implausible and disrespectful to insist that those who are oppressed are never culpable. Second, the belief that oppression *always* or even frequently exculpates the oppressor (in virtue of the oppressor's moral ignorance) seems to have the consequence that if one is potentially enjoying the benefits of oppressive social arrangements, on pain of acquiring new culpability one has reason to be unreflective about social arrangements.

Here is the core idea in what follows: oppression alters the morally adequate and deliberatively significant possibilities available to agents. Moreover, the way oppression interacts with the norms of blameworthiness (and the underlying requirements on culpability) explains a good deal of what is otherwise puzzling about culpability for wrongdoing under conditions of oppression. Given a context and a set of justified norms for that context, wrongful actions are culpable when the agent fails to exercise her moral considerations-sensitive capacities in ways that constitute a suitable quality of will.

Oppression can affect both what an agent can *do* but also an agent's *deliberations* and *beliefs*. Start with the case of oppression affecting what an agent can do. For an agent to lack due moral concern—to have ill will—the agent has to have available to her morally adequate possibilities of action. However, oppression can eliminate or reduce the morally adequate actional possibilities available to an agent. When the secret police require you to accuse your neighbor of disloyalty to the party, lest they kill or imprison your family, your accusing your neighbor has a different quality of will than when you accuse your neighbor with no corresponding threat. That is, one may act wrongly without a correspondingly poor quality of will.

Importantly, the fact of oppression does not always make agents acting under oppression not blameworthy, because oppression does not always remove all the morally adequate actional possibilities. So long as morally adequate options remain intact, and the agent remains a responsible agent (i.e., capable of discerning and suitably responding to the relevant moral considerations in that context), the agent can fail to have an adequate quality of will, even when subject to oppression.<sup>14</sup>

---

<sup>14</sup> In virtue of removing actional possibilities, oppression can count as a form of domination, as understood by Pettit (2015). Action under domination can be culpable if one retains responsible agency and has available morally adequate possibilities. This picture is compatible with a range of views about the relevance of the number of

Oppression can also alter epistemic agency by changing which possibilities are deliberatively accessible to the agent. If one has internalized oppressive norms or cultural scripts, these things shape one's deliberations about what actional possibilities are relevant in a given circumstance. In shaping an agent's deliberations, an agent may lose the relevant degree of access to deliberatively significant and morally adequate action possibilities, in comparison to circumstances devoid of that oppression. One effective way of perpetuating oppressive social conditions is for the costs of oppressive arrangements to be made obscure from those who benefit from those policies, whether by patterns of residence, schooling, media choice, or the formulation of policy. For members of advantaged groups, conditions of oppression might reduce the salience of, for example, complaints about aggressive policing in a racially marginalized community. Or, it might obscure the effects of particular economic policies on low-income wage earners. In general, privilege tends to obscure its costs imposed on non-privileged populations. If the loss of access to deliberatively significant and morally adequate action possibilities is sufficiently robust—that is, if in a large enough proportion of deliberatively similar circumstances the agent fails to recognize and respond to the relevant moral consideration(s)—then the agent will not be a responsible agent, and thus not culpable, at least with respect to those considerations in that context.

Not all weakenings of a capacity constitute an incapacitation. The loss of deliberatively relevant possibilities (read: the diminution of the proportion of relevant worlds in which the agent recognizes and suitably responds to the consideration) brought on by oppression may be only mitigating. An agent can satisfy the minimum standards for responsible agency, but the fact of oppression may also impair that capacity, blocking full culpability for wrongdoing. The wrongdoing by an agent impaired in that way does not express the same poor quality of will were the wrongdoing to have been performed by an unimpaired agent.

---

discernible morally adequate options for evaluation of an agent's quality of will. Presumably, an agent's failure of due concern is worse when there are a variety of readily discernible morally adequate options, as opposed to a case in which there is only one option that is difficult to discern. Moreover, as Christopher Lewis (forthcoming) has emphasized in a different context, *how* options are incentivized is plausibly relevant, quite apart from concerns about domination and injustice. Adaptive preferences and morally suspect socialization may work by foreclosing and incentivizing certain options, so that even if agents retain freedom in some or another sense, they can suffer from impairments to the sort of freedom required for moral responsibility. For a pertinent discussion, see Stoljar (2014).

Oppression can also affect the capacities for responsible agency without any material significance for culpability. Some instances of oppression can reduce the proportion of worlds in which the considered agent suitably recognizes and responds to the relevant moral considerations in the relevant circumstances, but where this reduction is insufficient to pull the agent's capacity underneath what is required for full culpability. Such changes of the agent's modal profile do not mitigate culpability. Moreover, sometimes oppression is entirely tangential to an agent's ability to recognize and suitably respond to moral considerations. In those cases, the fact of oppression does not affect culpability or blameworthiness. In these cases, the badness of oppression is not a function of its effects on the culpability of agents. Finally, oppression may *enable* agents to have a greater sensitivity to a relevant class of moral considerations, for example, in making certain kinds of injustice more visible to those who suffer from it. Oppression may thus enable responsible agency.<sup>15</sup>

The narrowness of the scope of exemptions, or the idea that someone isn't a responsible agent, is a distinctive feature of this account. If we think of responsible agency as requiring circumstantial sensitivity to considerations, typical losses of responsible agency are frequently patchy or variable across contexts. That is, we ordinarily lose responsible agency with respect to a *kind* of consideration in a *kind* of circumstance. Except in the most dramatic circumstances, one does not lose responsible agency with respect to all moral considerations across all possible circumstances. So it goes for losses of culpability due to oppression. The loss of responsible agency is typically neither total with respect to a class of considerations, nor global with respect to contexts. Nor are losses of responsible agency necessarily losses of other forms of agency—potentially, one may have diminished or absent capacities relevant to responsibility without necessarily having diminished agency with respect to other normatively or psychologically important conceptions of agency.

This approach also allows us to acknowledge some of the moral significance of oppression. Oppression can, and frequently does entail a kind of loss of agency. To the extent to which there are second- or third-personal reasons to want other agents to be competent at navigating moral demands, a loss of this form of agency is plausibly a loss of something we value.

---

<sup>15</sup> There is a closely related idea in standpoint epistemology about the way in which social position, including oppressed social positions, enables particular forms of knowledge. For an instructive discussion, see the introduction and first two chapters of Medina (2012).

To the extent to which we have first-personal reasons to want to be competent at navigating moral demands, there is a loss there, too. In either case, we might think such agency is inherently valuable, quite apart from an agent's actual regard for the possession of that form of agency.

Especially under conditions of unjust social inequality, the appropriateness of blame can have an asymmetric structure. For example, those who are socially disadvantaged may seem to have standing to blame those with social advantage, even when the reverse is not true, if those with social advantage are complicit in the creation or perpetuation of the social disadvantage. Oppressors can enjoy an especially nefarious social arrangement, blaming and punishing parties whose culpability is partly a byproduct of the choices of those with social advantage. Those complicit in the creation or persistence of oppressive social schemes plausibly lose their standing to blame the oppressed in those contexts (cf. Tadros 2009).

The extent to which oppressors are complicit in oppression that produces wrongdoing is often unclear. Disentangling real world causal and normative relations is seldom easy, and it is made even harder by the subtle ways oppression alters the epistemic agency and actional possibilities of agents. So, assessing culpability for wrongdoing under real world conditions of oppression is oftentimes an extraordinarily difficult task. The present approach permits us to make a few general observations, however.

First, even if concerns about complicity can affect the standing of the oppressors to blame the oppressed, blame will frequently be entirely appropriate from oppressed to oppressor. The former slave owner may lack standing to condemn the wrongdoing of the former slave, but this in no way precludes the slave from condemning the wrongdoing of the slave owner. Indeed, by the lights of the oppressed, those who are socially advantaged and permit the ongoing persistence of oppression will often seem negligent or even exceptionally ill-willed. Yet the facts about culpability are not a matter of perceived negligence. As noted above, oppressors can be insufficiently able to discern the relevant considerations, and thus, not culpable for their wrongdoing. So, many of the real-world difficulties in discerning culpability can be explained, but they are not readily eliminated.

There is a further dimension worth noting here. The responsibility system earns its keep in virtue of the role that moralized praise and blame play, over time, for creatures with psychologies like ours. Many of the psychological benefits are not at all direct, and it is an error to search for those benefits in every individual instance of holding one another responsible.

However, one underappreciated role of blameworthiness and its intimate connection to angry emotions of reaction is the work they do in animating subordinate groups to fight against conditions of oppression. Moral anger helps call attention to the wrong, but it can also bind people together in shared commitments to change social practices and institutions. The protestive dimension of moral blame can be a powerful tool in motivating and sustaining the kinds of coalitions that reshape social norms and the attendant moral practices (Smith 2013). Given that the primary normative function fulfilled by a system of blame just is the enhancement of moral considerations-sensitive agency, there is good reason to think that the norms of blame will typically support angry blaming by subordinate groups against superordinate groups.<sup>16</sup>

Let's take stock. Culpability under oppression is a function of some familiar features of our agency and the context-specific norms of responsibility as well as the way in which oppression interacts with those norms and the morally adequate options available to oppressed agents. One's status as a responsible agent or not, and whether one is blameworthy or not, are not facts prior to, or independent of the normative concerns and particulars of circumstance that structure our agency. Social context, in the form of individual or structural oppression, can impair or even undermine responsible agency. Other times, it leaves it wholly intact. The overriding lesson, though, is the relatively piecemeal nature of these interactions. Culpability is local, but variegated in how it arises from the matrix of agency and socio-normative forces.

#### **4. Moral ecology**

I have argued that agency is at least partly socially constituted, a product of an ecology that variously scaffolds and hinders morally appealing forms of agency. We rely on others to attune and calibrate our tracking of moral considerations (cf. Bennett 2002; McKenna 2012, 68; McGeer 2012; Vargas 2013a, 261-265; Hurley 2013). Social feedback and contextual cues do considerable work in norm inculcation, and failing that, in directing our attentions to what reasons there are in a given context. When the operative norms are not justified, practices of praising and blaming can reify oppression and other forms of moral error. If we accept that

---

<sup>16</sup> There is a complexity here having to do with the disposition of dominant groups to regard anger by those in oppressive circumstances as misplaced or unreasonable. I trust that concrete examples in contemporary political life are not hard to find. This significance of this phenomenon for the shape of a normatively ideal system of angry blame depends, in part, on whether there are non-instrumental justifications for moral anger.

culpability turns on whether we have particular capacities that are both products of and occurrently structured by social context, circumstances should not be an afterthought for a theory of moral responsibility. Rather, they are a central constituent of what it means for us to be responsible agents. So, the *moral ecology* of our agency matters precisely because it is one of the constituents of our responsible agency.

According to the agency cultivation model, the primary engine for the fostering of our moral considerations-sensitive agency is the internalization of moral norms and their consequent effects, especially in connection with culpability assessments. When moral norms are internalized, they structure how agents perceive actional possibilities, and they also provide an impetus for agents to attend to the relevant moral considerations. Our interest in being regarded as competent at navigating moral demands and avoiding blame shapes our prospective behavior. Similarly, we respond to others on the defeasible presumption that those agents are normatively competent and capable of deserving blame in light of how they exercise their agency.

In order to be effective, culpability norms must have sufficient currency in the relevant group, and this ordinarily involves common knowledge about the content of the norms.<sup>17</sup> However, social norms do not always require internalization to succeed in shaping the sensibility and conduct of agents. Awareness of a norm, especially combined with incentives (e.g., avoiding loss of status, avoiding unpleasant experiences of being blamed) can prospectively shape one's behavior, in ways often compliant but sometimes not compliant with the norm. Repeated norm compliance sometimes suffices to produce internalization of the norm. Sometimes it doesn't. With sufficient acceptance, however, prevalent norms can structure deliberation and conduct, orienting agents to the relevant moral considerations even when intrinsic motivation is absent.

Given a picture of agency that allows for its social constitution, we should expect that the content of laws, the shape of economic opportunities, and the various social scripts with currency will each structure our agency in diverse and oftentimes subtle ways. In addition to adopting and promulgating social norms with particular content, we might also hope to shape the moral ecology by attending to the way in which some of these other elements shape our agency. I conclude with some brief remarks about how a concern for culpability and fighting oppression

---

<sup>17</sup> Allowing for important variations between these accounts, one can find strands of these ideas in Copp (1995), Bicchieri (2006), McTernan (2013), and others.



might go if we take seriously the idea that we might alter our agency by altering the moral ecology of that agency.

From the outset, it seems clear that some interventions into our moral ecology—and thus, how we scaffold our agency—will be more and less palatable. Adjustments to our moral ecology that operate on our agency in ways that allow for a kind of rational mediation are preferable to interventions that bypass or subvert our sensitivity to moral considerations. It is now a familiar point that some changes to institutional “choice architecture” need not bypass considerations-sensitive agency (cf. Thaler and Sunstein 2009). Some so-called “nudges” can be interventions in a moral ecology that work through rational mediation. Others operate by exploiting fundamentally irrational processes, or otherwise bypassing an agent’s sensitivity to the relevant moral considerations.<sup>18</sup>

Some structuring of choices and environment seem compatible with rational mediation by the agent, and this is clear in non-moral cases. Putting salad rather than strawberry shortcake at the front of the lunch buffet—an intervention endorsed by Thaler and Sunstein—does not have to bypass someone’s rational agency. If the way this nudge functions is to permit the reasons to eat salad to be more salient, or more difficult to ignore, or less likely to be forgotten, there is little reason to think it subverts an agent’s rationality. So, if some interventions on moral ecology function in this way—by making moral considerations salient—there is less reason for qualms about those interventions.

Interventions that attempt to restructure local dispositions in ways that are not plausibly mediated by rational sensitivities are, on the face of it, less ethically palatable. Even so, once we accept a socially structured conception of agency, it is not always clear how we should think about mediation by rational agency. Conscious, reflective agency is relatively rare, and anyway, there is good reason to think it too is structured by situational and social pressures (Doris 2015). For accounts that emphasize that sensitivity to moral considerations can operate *de re*—i.e., without agents consciously thinking of them as such (e.g., Arpaly 2003)—it may be tenable to maintain that some interventions into our moral ecology will work in ways that are consonant

---

<sup>18</sup> There is some debate about how to characterize nudges. Some ways of regimenting Thaler and Sunstein’s construal of nudges makes their non-rational nature a matter of definition (e.g., Hausman and Welch 2010). On such construals of nudging, the present point is that there are many non-nudging adjustments to choice architecture that needn’t entail a bypassing of rational capacities.

with accurate informed reflection, but which may not require it in a given instance. Some mental processes are sensitive to considerations without active, conscious, deliberative assessment—e.g., reflexively pulling one’s hand away from a hot stove is rational in a range of contexts. So long as those processes are ones that reason can respect, as it were, the absence of conscious, deliberative moral assessment is less problematic (cf. Doris 2015, 52).

Private and individual adjustments to our ecology of agency are all good and well. However, by omission or commission, oppression is frequently a product of the social political organization of states. Even if we assume states will regularly tolerate some oppression, intentional attempts to shape circumstances to support responsible agency, and thus, culpability, might seem to require state intervention. However, given an antecedent commitment to some or another form of political liberalism, the possibility of direct state intervention into our moral ecology is not unproblematic. If states are prepared to exercise their power to back particular kinds of interventions, there is a traditional family of worries that states would be taking a de facto stance on substantive moral doctrines about which there is reasonable dispute. In doing so, some have worries that this is a move towards totalitarianism (Berlin 1958), or less dramatically, towards a kind of domination at odds with a serious commitment to individual liberty (Pettit 2015).

Four potential lines of reply are worth noting. First, liberal states *can* have grounds for building ecologies that favor agents being sensitive to some moral considerations and not others. For example, substantive moral considerations that are part of an overlapping consensus of what morality favors can be permissible under some forms of liberalism (e.g., considerations against genocide and slavery, and considerations grounded in widely-accepted human rights). More generally, in the absence of a substantive theory of morality and politics, reasonable constraints on such interventions may be readily identifiable. Such elements may include: requirements of publicity (i.e., that the interventions be known); respect for the moral, rational, and political agency of those affected by the interventions; an assessment of the net costs and benefits for various other moral and political ideals (e.g., autonomy, non-domination, dignity); and ongoing support and concern for ordinary forms of rational persuasion (cf. Hausman and Welch 2010; Hurley 2013).

Second, shaping the moral ecology, even to overcome oppression, need not require state intervention.<sup>19</sup> Individuals, institutions, and communities can make changes in the currency of social norms and substantially alter the scaffolding of our agency. Anti-bigotry interventions, from Stetson Kennedy's 1940s "Frown Power" campaign (which involved openly frowning at expressions of racism) to modern practices of "calling out" biased behavior, have re-shaped local practices. Community protests of businesses, norms of crossing picket lines, and boycotts are old and familiar tools in the fight against perceived injustice. Decisions about curriculum in schools, the promulgation of new workplace norms, and the framing choices as serving the public good may also contribute to combating oppression's culpability-undermining effects. Even if we accept the restriction on state neutrality about comprehensive moral issues, we still have important tools at our disposal to reshape the moral ecology. As Mackenzie (2008) argues in a different context, "ruling out coercive political means ...does not entail ruling out other political means for encouraging citizens to pursue valuable goals—for example, incentive and reward schemes" (529). Incentives and reward schemes are not tools limited to government actors, so we are not bereft of non-governmental resources for shaping our moral ecology.<sup>20</sup>

It remains an open question whether or not incentives, rewards, and limiting the reshaping of our moral ecology to methods and commitments palatable under political liberalism will provide for much in the way of ecology-shaping. In a posthumously published paper, Susan Hurley (2013) explores an interesting further option. She argues that traditional forms of political liberalism rely on a picture of agency that is at odds with our situated rationality. What is called for, she claims, is a revisionist and ecologically rational liberal conception of government. As she puts it, "Traditional liberalism tends to view government as the primary source of interference with responsible action by private individuals. [Her alternative, revisionist conception of]

---

<sup>19</sup> There is an interesting question about when deployment of state power is counterproductive. My suspicion is that alienation from the state or conditions of low social trust can undermine the effectiveness and credibility of state interventions.

<sup>20</sup> A caveat: with sufficient comprehensiveness in scope and distribution, as well as sufficient motivational force attached to incentives, a system of incentives can become a de facto system of coercion. If an enormous range of options are incentivized such that only a few options do not receive the benefit, those options will be perceived as costly. Such a system of incentives is coercive in all but name. This suggests that there is a conceptual problem for one way of thinking about domination and the constraints on republican government. However, for present purposes we only need the relatively innocuous thought there are ways to encourage norms and adjustments to choice architecture that don't amount to coercion.

liberalism should instead view government as having a counter-manipulative and positive role in the public ecology of responsibility” (2013, 209). On her account, a revised liberal form of government would be distinguished by a commitment to two principles that reflect a more naturalistically plausible conception of our agency. The first is a broadly negative principle, according to which government should regulate and counter manipulative influences on citizens’ beliefs and behavior. The second is a broadly positive principle, in which government is tasked with designing and creating a better public ecology for our rational agency. This latter project coheres in various ways with the project of nudges espoused by Thaler and Sunstein (2009), with a special emphasis on avoiding manipulative uses of government influences. So, re-envisioning the permissible scope liberal state interventions in a way that allows some non-trivial re-shaping of the moral ecology is a third possibility.

A final possibility is to simply reject the ideal of liberal neutrality, and to defend some or another alternative account for the basis of state (or other interventions) in the moral ecology.

We need not take a stand on which of these approaches, or what combination of them, is the most promising in fostering responsible agency under oppression. In supporting responsible agency and mitigating the effects of oppression, we do well to fashion all the tools we can.<sup>21</sup>

---

<sup>21</sup> My thanks to the Gunnar Björnsson, Bryce Huebner, Sofia Jeppsson, Per-Erik Milam, Ben Matheson, Sam Murray, David Pizarro, Grant Rozeboom, Tamler Sommers, Andras Szigeti, the members of seminars at UC Berkeley, University of Gothenburg, and Georgetown, as well as the terrific students and faculty at California State University Bakersfield, for discussions of ideas in this paper. I’m particularly grateful to Michael McKenna, Emily McTernan, and the editors of this volume for extensive written feedback on earlier versions of this paper.

## References

- Arneson, Richard J. 2003. "The Smart Theory of Moral Responsibility and Desert." In *Desert and Justice*, edited by Serena Olsaretti, 233–58. Oxford: Oxford.
- Arpaly, Nomy. 2003. *Unprincipled Virtue*. New York: Oxford.
- Bennett, Christopher. 2002. "The Varieties of Retributive Experience." *The Philosophical Quarterly* 52 (207): 145–63.
- Berlin, Isaiah. 1958. *Two Concepts of Liberty*. Oxford: Clarendon Press.
- Bichierri, Christina. 2006. *The Grammar of Society: The Nature and Dynamics of Social Norms*. New York: Cambridge University Press.
- Bierria, Alisa. 2014. "Missing in Action: Violence, Power, and Discerning Agency." *Hypatia* 29 (1): 129–45.
- Björnsson, Gunnar, and Karl Persson. 2013. "A Unified Empirical Account of Responsibility Judgments." *Philosophy and Phenomenological Research* 87 (3): 611–39.
- Brink, David O., and Dana Nelkin. 2013. "Fairness and the Architecture of Responsibility." *Oxford Studies in Agency and Responsibility* 1 284–314.
- Calhoun, Cheshire. 1989. "Responsibility and Reproach." *Ethics* 99 (2): 389–406.
- Cudd, Ann E. 2006. *Analyzing Oppression*. New York: Oxford University Press.
- Copp, David. 1995. *Morality, Normativity, and Society*. New York: Oxford University Press.
- Doris, John. 2002. *Lack of Character*. New York: Cambridge University Press.
- Doris, John. 2015. *Talking to Ourselves: Reflection, Skepticism, and Agency*. New York: Oxford University Press.
- Greene, Joshua and Jonathan Cohen. 2004. "For the law, neuroscience changes everything and nothing." *Philosophical Transactions of the Royal Society of London B*. 359:1775-1785.
- Hacking, Ian. 1995. "The Looping Effects of Human Kinds." In *Causal Cognition: A Multi-Disciplinary Debate*, edited by Dan Sperber, David Premack, and Ann James Premack, 351–83.
- Haslanger, Sally. 2012. "Oppressions: Racial and Other." In *Resisting Reality: Social Construction and Social Critique*, 311–38. New York: Oxford University Press.
- Haslanger, Sally, Tuana, Nancy and O'Connor, Peg. 2012. "Topics in Feminism", *The Stanford Encyclopedia of Philosophy* (Fall 2015 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2015/entries/feminism-topics/>.
- Hausman, Daniel M., and Brynn Welch. 2010. "Debate: To Nudge Or Not to Nudge." *The Journal of Political Philosophy* 18 (1): 123–36.
- Huebner, Bryce. 2016. "Implicit Bias, Reinforcement Learning, and Scaffolded Moral Cognition." In *Implicit Bias and Philosophy: Volume 1: Metaphysics and Epistemology*, edited by Michael Brownstein, and Jennifer Saul, 47–79.
- Hurley, Susan. 2013. "The Public Ecology of Responsibility." In *Responsibility and Distributive Justice*, edited by Carl Knight, and Zofia Stemplowska, 188–215. Oxford, U.K.: Oxford University Press.
- Levy, Neil. 2014. *Consciousness and Moral Responsibility*. New York: Oxford University Press.
- Lewis, Christopher. Forthcoming. "Inequality, Incentives, Criminality, and Blame." *Legal Theory*.
- Mackenzie, Catriona. 2008. "Relational Autonomy, Normative Authority and Perfectionism." *Journal of Social Philosophy* 39 (4): 512–33.
- Mackenzie, Catriona, and Natalie Stoljar. 2000. "Introduction: Refiguring Autonomy." In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, edited by Catriona Mackenzie, and Natalie Stoljar, New York: Oxford University Press.
- McGeer, Victoria. 2012. "Co-Reactive Attitudes and the Making of Moral Community." In *Emotions, Imagination, and Moral Reasoning*, edited by R. Langdon, and C. Mackenzie, 299–326. New York: Psychology Press.

- McGeer, Victoria. 2014. "P.F. Strawson's Consequentialism." In *Oxford Studies in Agency and Responsibility Volume 2*, edited by David Shoemaker, and Neal A. Tognazzini, 64–92. New York: Oxford University Press.
- McGeer, Victoria. 2015. "Building a Better Theory of Responsibility." *Philosophical Studies* 172 (10): 2635–49.
- McKenna, Michael. 2012. *Conversation and Responsibility*. New York: Oxford University Press.
- McKenna, Michael and Brandon Warmke. Forthcoming. "Does Situationism Threaten Free Will and Moral Responsibility?" *Journal of Moral Philosophy*.
- McTernan, Emily. 2013. "How to Make Citizens Behave: Social Psychology, Liberal Virtues, and Social Norms." *The Journal of Political Philosophy* 22 (1): 84–104.
- Medina, José. 2013. *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imaginations*. Oxford, U.K.: Oxford University Press.
- Mele, Alfred. 2009. *Effective Intentions: The Power of the Conscious Will*. New York: Oxford University Press.
- Nagel, Thomas. 1979. "Moral Luck." In *Mortal Questions*, 24–38. Cambridge, U.K.: Cambridge University Press.
- Nahmias, Eddy. 2007. "Autonomous Agency and Social Psychology." In *Cartographies of the Mind: Philosophy and Psychology in Intersection*, edited by Massimo Marraffa, Mario De Caro, and Francesco Ferretti, 169–85. Berlin: Springer.
- Nahmias, Eddy. 2010. "Scientific Challenges to Free Will." In *A Companion to the Philosophy of Action*, edited by Timothy O'Connor, and Constantine Sandis, 345–56. Malden, MA: Wiley-Blackwell.
- Nelkin, Dana. 2005. "Freedom, Responsibility, and the Challenge of Situationism." *Midwest Studies in Philosophy* 29 (1): 181–206.
- Nichols, Shaun. 2015. *Bound: Essays on Free Will and Responsibility*. New York: Oxford University Press.
- Oshana, Marina. 2014. "A Commitment to Autonomy is a Commitment to Feminism." In *Autonomy, Oppression, and Gender*, edited by Andrea Veltman, and Mark Piper, 141–62. New York: Oxford University Press.
- Oshana, Marina. 2015. "Is Social-Relational Autonomy a Plausible Ideal?" In *Personal Autonomy and Social Oppression: Philosophical Perspectives*, edited by Marina Oshana, 3–24. New York: Taylor and Francis.
- Pereboom, Derk. 2001. *Living Without Free Will*. Cambridge: Cambridge University Press.
- Pettit, Philip. 2015. "Freedom: Psychological, Ethical, and Political." *Critical Review of International Social and Political Philosophy* 18 (4): 375–89.
- Raz, Joseph. 2011. *From Normativity to Responsibility*. Oxford: Oxford University Press.
- Schmidt, Andreas. In progress. "Consequentialism and the Ethics of Blame." Unpublished manuscript.
- Smith, Angela. 2013. "Moral Blame and Moral Protest." In *Blame: Its Nature and Norms*, edited by D. Justin Coates, and Neal A. Tognazzini, 27–48.
- Stoljar, Natalie. 2014. "Autonomy and Adaptive Preference Formation." In *Autonomy, Oppression, and Gender*, edited by Andrea Veltman, and Mark Piper, 227–254. New York: Oxford University Press.
- Tadros, Victor. 2009. "Poverty and Criminal Responsibility." *Journal of Value Inquiry* 43 (3): 391–413.
- Thaler, Richard H. and Cass R. Sunstein. 2009. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New York: Penguin.
- Vargas, Manuel. 2004. "Libertarianism and Skepticism About Free Will: Some Arguments Against Both." *Philosophical Topics* 32 (1&2): 403–26.
- Vargas, Manuel. 2008. "Moral Influence, Moral Responsibility." In *Essays on Free Will and Moral Responsibility*, edited by Nick Trakakis, and Daniel Cohen, 90–122. Newcastle, UK: Cambridge Scholars Press.
- Vargas, Manuel. 2011. "The Revisionist Turn: Reflection on the Recent History of Work on Free Will." In *New Waves in the Philosophy of Action*, edited by Jesus Aguilar, Andrei Buckareff, and Keith Frankish, 143–72. New York: Palgrave Macmillan.

- Vargas, Manuel. 2013a. *Building Better Beings: A Theory of Moral Responsibility*. Oxford, U.K.: Oxford University Press.
- Vargas, Manuel. 2013b. "Situationism and Moral Responsibility: Free Will in Fragments." In *Decomposing the Will*, edited by Till Vierkant, Julian Kiverstein, and Andy Clark, 325–49. New York: Oxford University Press.
- Vargas, Manuel. 2016. "The Runeberg Problem: Theism, Libertarianism, and Motivated Reasoning." In *Free Will and Theism: Connections, Contingencies, and Concerns*, edited by Kevin Timpe, and Daniel Speak, 27–47. New York: Oxford University Press.
- Watson, Gary. 1987. "Responsibility and the Limits of Evil." In *Responsibility, Character, and the Emotions*, edited by Ferdinand David Schoeman, 256–86. New York: Cambridge.
- Westlund, Andrea C. 2009. "Rethinking Relational Autonomy." *Hypatia* 24 (4): 26–49.
- Wolf, Susan. 1987. "Sanity and the Metaphysics of Responsibility." In *Responsibility, Character, and the Emotions*, edited by Ferdinand David Schoeman, 46–62. New York: Cambridge.
- Young, Iris Marion. 2004. "Five Faces of Oppression." In *Oppression, Privilege, and Resistance*, edited by Lisa Heldke, and Peg O'Connor, 37–63. Boston, MA: McGraw Hill.
- Zawidzki, Tadeusz. 2013. *Mindshaping: A New Framework for Understanding Human Social Cognition*. Cambridge, MA: MIT Press.