

Marquette University

e-Publications@Marquette

Dissertations (1934 -)

Dissertations, Theses, and Professional
Projects

Looking Through Whiteness: Objectivity, Racism, Method, and Responsibility

Philip Mack
Marquette University

Follow this and additional works at: https://epublications.marquette.edu/dissertations_mu



Part of the [Philosophy Commons](#)

Recommended Citation

Mack, Philip, "Looking Through Whiteness: Objectivity, Racism, Method, and Responsibility" (2022).

Dissertations (1934 -). 1200.

https://epublications.marquette.edu/dissertations_mu/1200

LOOKING THROUGH WHITENESS:
OBJECTIVITY, RACISM, METHOD,
AND RESPONSIBILITY

by

Philip Thomas Langlois Mack, A.B., M.A.

A Dissertation submitted to the Faculty of the Graduate School,
Marquette University,
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy

Milwaukee, Wisconsin

May 2022

ABSTRACT
LOOKING THROUGH WHITENESS:
OBJECTIVITY, RACISM, METHOD,
AND RESPONSIBILITY

Philip Thomas Langlois Mack, A.B., M.A.

Marquette University, 2022

Does a white philosopher have anything of value to offer to the philosophy of race and racism? If this philosophical subfield must embrace subjective experience, why should we value the perspective of white philosophers whose racial identity is often occluded by racial normativity and who lack substantive experiences of being on the receiving end of racism? Further, if we should be committed to experience, in what sense can the philosophy of race and racism be “objective”? What should that word mean?

Tackling this question first, “objective” should at least mean *general*, that the ideas of the literature can be coherently integrated. An objective take on racism brings together a plurality of perspectives. What’s wrong with *just* a plurality of satellite ideas? It implies a fragmented approach to ameliorating racism, where different specialists have different recommendations. How can racism, generally, be lessened? If major views of racism are unifiable, then we have a general method to ameliorate racism.

This project might appear tone-deaf: a white philosopher unifying things by reducing ideas to some central notion. But this unity isn’t about reducing things but rather integrating them in a way that respects difference. Yet, there’s a reason we should be interested in the white perspective. Whites can speak about racism from a *participatory perspective*. If whites are knowledgeable, and believe themselves to have no implicit bias, they may suppose they’re “beyond” racism or no longer at risk for perpetuating it. I explore this idea in a psychologically realistic way via my notion of *overlooking*, where ameliorating racism from the white perspective is an ongoing project.

I end by considering how racism is applicable to other philosophical ideas beyond its typical or circumscribed purview. Here, I re-frame responsibility, arguing that we needn’t be forced to choose between responsibility models divided into individual versus social camps. We ought to instead think of responsibility in terms of power, which provides a realistic lens by which persons and groups are held to account. In being more generally convincing, it might actually get folks to take responsibility where they might not otherwise — theory in service of praxis.

ACKNOWLEDGEMENTS

Philip Thomas Langlois Mack, A.B., M.A.

Grant Silva's support and reassurance afforded me the confidence to pursue this project. He read every word – drafts to final product. His feedback, conversation, and challenges helped me refine my ideas, strengthen my positions, and unearth my philosophical identity. He treated me as a peer. I can't thank him enough.

Thank you to all my readers, Mike Monahan, Anthony Peressini, and Theresa Tobin, for your commitment to my project and for your patience. Most of all, thank you for your confidence in my work and for finding my ideas promising enough to join me on seeing this project through.

I owe many thanks to several friends whose conversation and guidance were formative, and whose support propped me up when I was falling down. Thank you to Clark Wolf, Shaila Wadhvani-Greenhalgh, Jorge Montiel, Kevin Gibson, Kiesha Martin, Sarah Kizuk, Greg Trotter, Bill Kollatz, and Raj Pelon.

I want to single out Pete Burgess for special thanks. His exceeding patience in helping me develop my views can't be overestimated. Although I couldn't meet all of his philosophical challenges in this dissertation, I'm reasonably sure that if I had, the work would have approximated the truth.

Finally, I owe considerable gratitude to my parents Lynn and Tom Mack who didn't bat an eye when I first decided to pursue philosophy and haven't since wavered in their support. Their encouragement, understanding, and patience sustained me through the entire process of writing this dissertation and well beyond.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	i
CHAPTER 1: WALKING TIGHTROPES: CHARTING A PATH REplete WITH TENSIONS.....	1
1. Whither Philosophy and the White Philosopher?	1
2. Contextualizing Tensions	6
3. Method	21
4. Previews.....	28
CHAPTER 2: RACISM, AFFILIATION, AMELIORATION: WHY NON-REDUCTIVE UNIFICATION MATTERS	33
Introduction.....	33
1. Putting Together the Picture	37
1.1. Horn 1: Selections of the “Many”	38
1.2. Horn 2: Reduction, or the “One”	44
1.3. Dissolution: Method and the Affiliation Diagnostic System.....	49
1.3.1. The Affiliation Diagnostic System	52
1.3.2. Benefits of the Diagnostic System	58
2. Unification to Amelioration	62
2.1. Varieties of Racism, Ways of Ameliorating	63
2.2. A Generalist’s Apology.....	69
Conclusion	72

CHAPTER 3: <i>OVERLOOKING: IT COULD HAPPEN TO YOU</i>	75
Introduction.....	75
1. The Overlooking Phenomenon.....	77
1.1. What is Overlooking?	81
2. Active Ignorance, Implicit Bias, and Overlooking: Continuities and Departures	82
2.1. Active Ignorance and Overlooking.....	83
2.2. Implicit Bias and Overlooking.....	90
3. Three Impacts of Overlooking	96
3.1. Description.....	97
3.2. Internal Regulation and Amelioration	100
3.3. Moral Evasions.....	102
Conclusion	106
CHAPTER 4: A CONTINUUM VIEW OF RESPONSIBILITY	109
Introduction.....	109
1. Surveying Options.....	111
1.1. The Individual.....	111
1.2. The Social	114
2. Responsibility on a Continuum.....	123
2.1. Continuum I: Harm.....	124
2.2. Continuum II: Power and Agency	129

2.3. Continuum III: Responsibility	135
2.4. Conscious Choice and Unawareness	139
Conclusion	141
CONCLUSION: CHARTING OUT NEW PATHS: TWO SITES FOR FUTURE RESEARCH.....	143
1. The Privilege of Responsibility	144
1.1. The Privilege of Taking Up Responsibility	145
1.2. The Privilege of Doing Good and of Praise	147
2. The White Philosopher and the Aggressor/Oppressor Perspective	148
2.1. The White Philosopher Approaches Whiteness and White Racism	149
2.2. Microaggressions from the Perspective of the Aggressor/Oppressor	150
BIBLIOGRAPHY	153

Chapter 1 Walking Tightropes: Charting a Path Replete with Tensions

...Sullivan discusses what we might call the white liberal double bind. This is the situation that well-intentioned white people find themselves in when they sincerely want to deal productively with race-related issues but seem condemned to screw up no matter what they do.

Paul C. Taylor (2007, 202)

Our craving for generality...[is also]...the contemptuous attitude towards the particular case.

Ludwig Wittgenstein (1958, 18)

1. Whither Philosophy and the White Philosopher?

Permit me to bombard you, reader, with a deluge of questions. Is it appropriate that white people participate in the philosophy of race and racism, especially given that the experience of racism is often missing for them? Can they *do* philosophy of race and racism? Perhaps so, perhaps not. Suppose it is appropriate. Suppose they *can* do it. What then, if anything, can a white philosopher (for instance, this author) have to say about issues the field discusses?

Bolder yet, can a white philosopher say anything *objective* about race and racism?¹ On its face, it would appear to be, *prima facie*, implausible that white philosophers have anything to offer the philosophy of race and racism seeing as there are good reasons to respect and think through particular subjective perspectives, especially nonwhite perspectives, where it comes to race and

¹ As I explain throughout this chapter, and in chapter 2, “objectivity” here means a *general*, common ground from which anyone can proceed and to which anyone can in principle agree. I will complicate and expand this notion soon enough.

racism.² So, should this project stop now before it goes any further? Am I doomed by that which is expressed in this chapter's first epigraph? Am I a pretender to the field, attempting to answer questions that I have at best no business investigating and at worst no way of sincerely entertaining? The burden is on white philosophers to justify how it is that their *particular* experience of whiteness or white identity is relevant to the production of philosophical knowledge pertaining to race *and* racism. This dissertation, in some ways, offers an initial foray into charting this ground. But equally important, who I am is relevant to writing *this* particular dissertation. It is made possible *because* I am white.

But what of philosophy itself? Of what value can philosophy be in this discussion? Does philosophy in its "traditional" sense – seeking universal, timeless truths – have anything to contribute to a conversation about race and racism? The local, historicized nature of race and racism would seem to preclude that entirely. This tension is grounded by Charles W. Mills' (2012) point that "traditional" philosophical inquiry's experiential starting point can be understood to be white (60).³ He writes that

...the conception of the discipline itself is inimical to the recognition of race. Philosophy is supposed to be abstracting away from the contingent, the corporeal, the temporal, the material, to get at necessary, spiritual, eternal, ideal truths. Because race as a topic is manifestly not one of those eternal truths...it is handicapped from the start.... Philosophy aspires to

² These reasons are briefly spelled out below and in more detail in chapter 2.

³ He notes that this is especially true of political philosophy (61-63, 65). Indeed, his point is that even in this sub-field of philosophy where we might expect to find an emphasis on the particulars of political life (e.g., as historically and locally contextualized), we find instead a striving for clean, abstract, "colorless" ideals (61).

the universal, whereas race is necessarily local, so that the unraced (whites) become the norm (60).

He later notes that

[w]ithout a willingness to face how seemingly colorless abstraction is really generalization from the white experience, the discipline's exclusions, both demographic and theoretical, can only perpetuate themselves (65).

If my experience as a white person is relevant to the investigation of race and racism, and yet I want to aim for a general philosophical approach to the field, then instead of abstracting away in a "colorless" fashion, I will need to sincerely confront, and not evade, the sorts of issues Mills sets forth.

So, is a "generalist" approach – putting together the various ideas in the literature clearly and coherently – therefore doomed in the face of issues that are particular by definition? How are white philosophers to confront the challenges Mills sets forth? Perhaps there is a middle ground where the fruits of particularity interweave with the descriptive simplicity of pointing out what is common. Couldn't it be coherent to crave generality *without* a contempt for the particular? If it is, might there be something we can learn about racism in general that isn't only about a narrow, specialized problem in the field – e.g., active ignorance, implicit bias, intentional hatred, bad faith, and so on?

The foregoing is in part the motivation for treating philosophy in this dissertation as an ameliorative activity in the spirit of American pragmatism. We need a plausible *path* forward, yet one that is attentive to differences at the level of the particular (e.g., experience, location, history, and so on). I understand

philosophy's ameliorative function as both *reflexive* and *practical*. By its "reflexive" function I mean that philosophical analysis should "look at" itself and self-correct for the purpose of avoiding its pitfalls. In the context of this dissertation, this unfolds both in terms of the white philosopher's self-analysis as a white philosopher – what we can do and what we ought to do when theorizing about race and racism – as well as whether philosophy itself can still aspire to something objective in the face of the localized nature of some of the problems it addresses (e.g., race and racism). Philosophy's ameliorative function as "practical" means that philosophical activity ought to attempt to make our lives better and that philosophers ought to act as functionaries of humanity.⁴

I raise the distinction between *reflexive* and *practical* functions for two reasons. First, it is meant to clarify that amelioration does not *solely* have to do with improving the conditions of real life. Part of my point is that philosophical theorizing can *itself* be made better for the purpose of addressing problems that it is, under certain conceptions, not particularly well-suited to address (as we have seen above vis-à-vis Mills).⁵ Second, in the context of racism, we should not be too quick to suggest that amelioration will "solve" the problem of racism. I take seriously Derrick Bell's (1992) "racism is permanent" thesis (373-374). Bell sets forth this thesis so as to combat the tendency of our thinking that racism can be

⁴ But see note 24 of this chapter where I complicate the notion of *practical* in juxtaposition to *praxis*. I expand on the notion of philosophers as "functionaries of humanity" in section 3.

⁵ Chapter 2 of this dissertation attempts to perform exactly this task by offering a meta-theoretical approach to analyses of racism in order to show that one need not clear away the contingencies and particularities of race and racism in order to be able to say something general, yet productive, about racism, thus confronting the challenge presented above from Mills.

overcome, as this tends to overdetermine plausible solutions to racism. Instead, Bell suggests that we find value in struggle and *not* in overcoming. This matters in the context of this project because I am not suggesting that amelioration requires or entails a complete solution to the problem of racism. We just need to do better.

The final consideration I want to raise by way of introduction is as follows. Beyond the domain of racism proper, suppose there is something to learn from racism as a general position – that is, as a *philosophical method*. Might it help our thinking about other intractable philosophical problems? If we gather up the lessons and insights derived from philosophical work on racism, we might be able to extend to other areas of philosophical thought more generally, using the problem of racism, in a sense, as an approach. As a social-historical phenomenon, racism betrays strict categorization into one of either the individual or the social/structural, and yet we presumably want to capture both for a full expression of the phenomenon. But capturing both levels of analysis extends far beyond the domain of racism to other areas where lines are blurred. Is making things right where racism is concerned a matter of changing the minds of racists, or is it a matter of upending the structures that reify racism? Does the one bear on the other? How? And how might the slipperiness of the individual/social-structural binary impact our thinking about the moral idea of responsibility, where structural and collective harms are not so neatly tied to individuals and vice versa?

2. Contextualizing Tensions

This section names and makes explicit the tensions just presented for the purpose of situating this chapter and the dissertation more generally. It charts a path replete with tensions which this dissertation, as if walking a tightrope, attempts to maneuver. The point is to introduce some overarching themes the dissertation addresses and to set the stage for a discussion of method—*how* to maneuver the tensions and address its themes.

Objectivity/Generality, Subjectivity/Particularity

It matters that we ask what the role of the white philosopher is in the philosophy of race and racism, and whether it is appropriate, because it is not straightforwardly clear that white people have the epistemic access that oppressed persons of color do regarding racism. Indeed, “not straightforwardly clear” may be putting it lightly. It is probably the case that whites *by and large* do not have that access, at least to the extent that persons of color do.⁶ It matters also that we ask whether a white philosopher can say anything *objective* because as a matter of method most of the field takes a perspectival approach, embracing experience. Both issues should be of interest to any white philosopher working in this area. We (white people) need to confront the issues honestly and carefully.

⁶ Of course, this is not to argue that white people do not experience class-based, sexual or gender, or other forms of oppression. It is to say, however, that the *experience of racial oppression is often missing*. I do not mean to overplay the lack of experience whites have on being on the receiving end of racism. Surely some have encountered some form of oppression and can relate in a way to racism as experienced by persons of color. A lifetime of accumulated experiences with some form of oppression likely develops a kind of empathy or moral imagination as concerns racism.

So, why is it that these problems run so deep? The white philosopher's legitimacy is challenged by a constellation of issues. Perhaps the most important intellectual predecessor here is W. E. B. Du Bois (1986) who argues that white minds are conditioned by "long followed habits" of white supremacy calcified in their environments (679). The notion of habit continues to play an important role in the philosophy of racism. Beyond their underscoring the maintenance of white supremacy and superiority (often unnoticed) habits of privilege and whiteness are thought to obscure what whites can know about race and racism (Sullivan 2006; MacMullan 2009). This white epistemic obscurity is drawn out by Mills' (1997) "epistemology of ignorance", the view that whites actively keep themselves "in the dark" about racial issues so as to maintain a racist status quo. If I keep myself from knowing, then there is nothing to see, and thus nothing for me to do. The payoff is that I get to keep all the privileges afforded to me by that (unjust) status quo. Worse yet, whites may not know that they do not know about relations of racial oppression (a so-called "meta-ignorance"), and in line with maintaining the status quo, they may want or need to keep it that way (Medina 2013, 35). This cluster of thought is amusingly captured by Terrance MacMullan (2015):

...when a white philosopher writes about whiteness, he or she says..."Hey! You should listen to what I have to say about race and whiteness" and then almost invariably explicitly states...that white folks suffer from deep ignorance when it comes to race and racism ("One of the things I have to say about race is that white people, like me, don't know what the heck we are saying when we talk about race!") (647).

Now, the more provocative element of this tension is that between a white philosopher's ability to say something *objective* about racism and the perspectivalism presupposed in much of the literature. *Prima facie*, there seems no way around this. "The white man," says Frantz Fanon (1952), "is locked in his whiteness" (xiii). If we are predominantly bound by our subjective perspectives, then there seems to be no clear route towards saying anything objective at all, where being objective means thinking from a detached, decontextualized standpoint. It is precisely the notion of *perspective*, of course, that is the challenge here. Whiteness is a perspective from which I understand the world around me (Frankenberg 1993, 1; hooks 1989, 113). But if this is so, and whites are often missing the experience of racial oppression, then it seems likely that whites cannot fully understand what racism *is like*, much less offer an objective take on it.

So, what is to be said about this often-missing experience and whether white philosophers can say something objective about race and racism, both *at all* and without homogenizing the discussion as warned against by Mills (2012)? Let's begin unpacking this by clarifying the sense of "experience" in this project before moving on to complicate the notion of "objectivity" in relation to subjectivity and particularity.

I use the term *perspective* (or "subjective perspective") to denote two kinds of experience: (1) *subjective experience* and (2) *experience-based knowledge*. Whereas (1) signifies one's first-person conscious experiences, (2) signifies an epistemic

feature of experience – the knowledge garnered from one’s experience. One’s subjective experience (1) will vary depending on one’s race. And one’s experience-based knowledge (2) will vary too, as it depends on the experiential differences based on one’s race. Thus, one’s experience-based knowledge is going to be *perspectival* – i.e., different to others’ experience-based knowledge. My point is that it is not *merely*, or even predominately, an *experiential* difference that is at stake here, but that it is an *epistemic* difference based on those experiential differences. So, all I mean by a *perspective* (or “subjective perspective”)⁷ is an experience-based knowledge.

Now, how is the abovementioned notion of perspective supposed to work in tandem with objectivity, especially where we do not want to erase raced-based perspectives? I propose carving out a middle ground where objectivity *and* raced-based perspectives are retainable. But doing so requires problematizing and complicating the notions of *objectivity*, *subjectivity*, and *particularity*.⁸ We may rightly want to avoid, methodologically, holding stalwartly to either subjectivity/particularity or a “traditional” sense of objectivity.⁹

To see why, let’s continue situating the discussion by re-raising the following question: what is the role of philosophy here? Given that the prevailing methodologies in the philosophy of race and racism belie generality, is

⁷ I will use the terms “perspective” and “subjective perspective” interchangeably.

⁸ Conspicuously absent from this list is the notion of *generality*, but the reader will soon find that generality is included in the sense of “objectivity” I offer in this section and below in section 3.

⁹ The “traditional” sense I have in mind is that against which Mills (2012, 60-65) warns – that it means a pure, “ideal” and timeless truth divorced from the messy particulars of the real world.

there room for a generalist—someone who wants to objectively integrate things—in the philosophy of race and racism? Can philosophy aspire to *anything* general in this area? We may not want to offer a full-throated generality, as it may come in tow with a simplistic reduction, the implication of which is doing away with the perspectives of philosophers of color who theorize from the oppressed perspective.

Now, although it would be inaccurate to say that all race theorists share the same methodology, I think it is fair to say that nearly all of them begin and end at particularity of some sort—an approach we can call *methodological particularism*. Sometimes that means concentrating on particularized experiences (Gordon 1995; Jones 2009, 31-33, 36; Alcoff 2012, 36; Mills 2012, 60-65; Kim 2014; MacMullan 2015, 646; Lee 2020). Elsewhere it amounts to a specified concentration on a particular time and location: an Anglo-American politic (Mills 1997, 2017) or legal history (López 2006), the historical residue of apartheid in South Africa (Vice 2010), vestiges of colonial whiteness as an educational norm and as “elite” in Nigeria (Ayling 2019), historical processes (Roediger 1991, 2002, Allen 2012a, 2012b; Yancy 2012, 5), or social location in the United States (Alcoff 2006, 9; Sullivan 2006, 3; Yancy 2008, xvi; MacMullan 2009, 1-2). I could go on, but one gets the point.

Now, there is nothing wrong with particularity *per se*. But what results from a singular commitment to particularity is *a proliferation of predominantly independent, satellite theories of racism*. The literature becomes fragmented,

obfuscating our understanding of racism, a consequence we might rightly want to avoid. But as noted before, we should not be too quick to dismiss particularity, as that runs the risk of doing away with first-person perspectives. It would be a mistake, for instance, to clean things up by offering a reductive theory of racism, that racism is, or boils down to, x , whatever “ x ” may be.

Yet, if subjectivity and particularity are embraced, then that likely implies a commitment to epistemic and moral relativism. That race and racism are so historicized and localized, it might make sense to suppose that we ought to theorize only from race-based perspectives. But if this is our only route, then there is no principled basis on which to say racism is morally objectionable. If it is all relative, then how is it that we can stake that claim? Clearly, if we cannot say racism is morally objectionable, then something has gone wrong with our account of racism (as against the background of subjectivity/particularity). More still, there would seem to be little to no common ground on which we can agree about the problems of race and racism, if it is true that we should theorize only from our perspectives.

But the traditional sense of “objectivity” does not do much better. Suppose we commit to the notion. It would likely follow, then, that no subjective perspectives matter at all. We may proceed thinking that raced experiences only get in the way of our arriving at the “Truth” with a capital “T”. So, we eliminate those perspectives, effectively sanitizing philosophical discourse of any “color”. But surely this is not the path to walk if philosophy of race and racism is our

focus. To pledge ourselves to this sense of objectivity is to pledge ourselves to philosophizing against the backdrop of white normativity – that “real” philosophy is colorless, that white is “right”. This would be to ignore perspective altogether. Thus, either subjectivity and particularity mean nothing, or there is nothing common to speak about. This is precisely the tightrope I seek to traverse.

So, what if instead of maintaining either of subjectivity/particularity or the traditional sense of “objectivity”, we massaged these notions to find a middle path? My proposal is to think of “objectivity” as a general common ground, something to which we could all in principle agree, where “general” is construed pluralistically, and the common ground achieved is an integrative plurality of perspectives. Note that this does not mean a so-called “view from nowhere”. Rather, it is a view to which anyone with different perspectives could in principle agree. The point is this: I (the author) do not completely know the perspectives of others; thus, I have to listen to others. But I cannot listen if I presume there is nothing we can agree on. Presuming so would be to trivialize other points of view. My point is that all perspectives are all equally different, and the only way to respect that difference is to assume we have common ground for agreement. We listen.

The “common ground” to which I refer, and the work my non-standard use of “objectivity” does for this project (in the context of race and racism), is constitute a view from which we add up the perspectives and particularities, and

we all see the same thing.¹⁰ Now, this does not mean that we are disembodied, or that we “leave” our perspectives, or that we think ahistorically. All it means is that we can all see things we can agree on, in principle, intersubjectively.¹¹ It is an objectivity that necessitates plurality and difference. It does not look at racism to find *sameness*, but rather, to find *difference*. It seeks out multiple accounts of what racism is, and how it can be modeled, because it is in those accounts that we strengthen the objective view as opposed to discounting difference so that it fits our preconceived notions of racism. Thus, nothing in my notion of objectivity should be taken to imply anything like a white normative theoretical standpoint, that objectivity means *whiteness à la* Mills (2012).¹²

Thus, instead of fighting against the blockades of ignorance and subjective perspective, I will attempt to put together a coherent picture of racism, without decontextualizing it and *while* retaining subjective perspectives. The positive consequence of doing so is our ability to speak generally about racism without denigrating particularity. It also has the benefit of our being able to think clearly about how the phenomena described by independent, satellite theories of racism may be ameliorated. If we are clear on what we are talking about, and clear on what we are picking out, then we can more clearly strategize how to make things better.

¹⁰ I am not suggesting here that we all need to have the *same* perspective to agree on something.

¹¹ The alternative is epistemic relativism – that none of us could possibly agree on anything.

¹² Neither should it be taken to imply that nonwhite philosophers cannot operate with my notion, the very suggestion of which would itself be racist.

So, perhaps there is yet a role for the white philosopher, and the possibility of speaking objectively about racism. What the white philosopher *can* do is an analysis *of* analyses of racism, the argument of chapter 2, and describe the experience of racism from the standpoint of the *participant*, the argument of chapter 3. The former is a second-order analysis which seeks no answer to the “What is racism?” question and ignores no first-person perspectives.¹³ The latter is a first-order description of what it is like to watch racism unfold (and often participate in the unfolding) from the perspective of a white person. These are the ways I answer whether the white philosopher can speak objectively about racism.¹⁴

Why the White Perspective Matters

That whites often lack the experience of racial oppression is one thing, but there is also a worry that they may do real damage if they go ahead theorizing from their perspective anyway. The worry is that white philosophers run the risk of “whitely” homogenizing the conversation. Whites who critique whiteness may “participate in the object of [their] critique” without realizing it by centering the conversation on whiteness *from* the white perspective (Ahmed 2007, 150, 158). If to theorize is to uncover the assumptions that often go into our thoughts, then

¹³ As the reader will find in chapter 2, I do not mean to suggest that this is *all* that the white philosopher can do, or indeed, the *only* approach available to white philosophers.

¹⁴ Now, lest I equivocate two senses of “objective” here, permit me to dispel any tension. Whereas in chapter 2, “objectivity” is achieved by taking a second-order look at the first-order landscape (i.e., theories of racism), in chapter 3, “objectivity” (though I do not rely explicitly on the term in that chapter) is understood from within the domain of white experience. So, the former is a domain anyone in principle can agree to because it attempts to assume no perspective in particular, and the latter is a domain where the description is true *within* the domain’s universe of discourse (i.e., the white subjective perspective).

it is white people uncovering whiteness – layers of the onion peeled back all the way through.

The implication is that critical approaches to race and racism may be whitewashed – erasing or ignoring the perspectives of the oppressed. This worry is further expressed by notions like “white solipsism” (Rich 1979, 299) and “ontological expansiveness” (Sullivan 2006, 10), an “impartial” “whiteness” (Frye 1992, Ch. 14), the “whitely” eye (Taylor 2016), and “loving, knowing ignorance” (Ortega 2006), all of which share a common concern: whites tend to think about the world from a privileged lens, rendering persons of color invisible. Thus, whites might engage in “white-splaining” to borrow a phrase from MacMullan (2015, 648), unwittingly stamping out or ignoring the insights of philosophers of color in the course of doing their work.

So, where does this leave white philosophers of race and racism? Should we circumvent the white perspective altogether to avoid the issues cited above? I worry that doing so would amount to leaving on the table how the white mind works and whether white racism might be mitigated. The white perspective should be *embraced*, we (white philosophers) need to *think through* the white perspective.

Whereas, thus far, a case has been developed for a level of generality for the analysis of race and racism, here the case pivots to the level of particularity. What does it feel like to participate in whiteness from a first-person perspective,

that is, from a position of power and privilege?¹⁵ Such an approach matters because if we know how white people think, then perhaps we can take that into account for the purpose of making things better – the ameliorative “thrust” of the approach. If we care at all about lessening racism, we need to reach whites who are not entirely closed off to acknowledging racism and their privileged position. That means setting forth a realistic white psychology that is not reductive or simplistic (e.g., that they should feel guilty, feel shame, or adopt a constant state of vigilance against possible racism).

Bluntly put, if ameliorating oppressive relations is a practical goal, don't concentrate on the oppressed and understand how *their* minds work, concentrate on the *oppressors* and how *their* minds work. The reason is simple: whites have privilege. And what is privilege but a kind of *power*? Understanding how that power operates in relations of oppression requires understanding how that power operates in white psychology so as to mitigate its possible harms. Just as a defensive coordinator in the NFL must study and know the opponents' offensive schemes and mindset, the white philosopher concerned with race and racism needs to study and know her own mind. The other teams may continue scoring points throughout a season and in seasons to come (racism is “permanent”), but good game plans mitigate how frequently points are scored (racism is ameliorable).

¹⁵ To be clear, this “pivot” is not precluded by what has so far been said – the particular and the subjective are not ruled out by objectivity or generality, but rather *retained*.

Integration and Dis-integration

So far, I have supposed that racism is ameliorable. Now, the strategies for ameliorating racism are kaleidoscopic. And rightly so. Racism is not *just* a matter of individuals behaving badly, but a broad matter of social structures continually churning out oppressive conditions. The nature of what is to be ameliorated oscillates from the individual, the inter-personal, and the structural. It is as if we are playing whack-a-mole, at one moment setting our sites on the individual, the next on the inter-personal, and then again on the structural, only to miss out on one or the others. That is, where we gain traction on strategizing how to ameliorate racism from the perspective of the individual, doing so may come at the expense of ameliorating racism at a structural or inter-personal level, and vice versa. So, the tension is this: how do we capture one without excluding the other? Why should the strategies be dis-integrated instead of *integrated*?¹⁶

Perhaps the tension arises from a question of which *kind* of amelioration strategy is more important, more valuable, more urgent, or more useful. While this is worth exploring in its own right, I will not touch it here. The point is that the literature often wavers in locating the site for making conditions better. Amelioration, therefore, becomes murky business. But the power of the generalist approach adumbrated above is precisely that getting clear on how to talk about racism enables getting clear on its mitigation. What this involves is an

¹⁶ I write “dis-integrated” with a hyphen instead of without, because I do not mean to imply that there is already an existing whole that has been broken up (as in *disintegrated*, *sans* hyphen). By “integrated” I just mean that complex parts can be put together into a unified whole.

attempt to bridge the divide between the individual and the structural, a theme that plays out over the course of chapters 2 and 3. Capturing both requires theoretically re-framing how to think about amelioration programs as they spring forth from certain descriptive commitments about the nature of what is to be solved – is it the hearts and minds of individuals, a community’s shared attitude, the institutions comprising society?

Though I draw this tension between integration and dis-integration from the literature on racism, it has relevance and broad impact in other areas of philosophy. Just as amelioration strategies suffer from blurred lines, the literature on responsibility shares a similar problem. Even a cursory look at the *Stanford Encyclopedia of Philosophy* entries on “Moral Responsibility” (Talbert 2019) and “Collective Responsibility” (Smiley 2017) reveals a clutter of metaphysical disputes. What makes an *individual* responsible? What makes it such that individuals constitutive of a group *share* responsibility? What makes a *collective*, independent of individuals, responsible?

Metaphysical morass ensues in answering these questions, and we end up with a murkiness similar to amelioration. Where the individual is responsible for actions she can reasonably author, she may not be responsible for her complicity in structural harms. Where individuals share responsibility for some harm, some may be unfairly implicated. And where a collective is responsible, some individuals may feel no personal responsibility at all. The point of raising all of

this is to suggest that we may want to re-frame how we think of the idea of responsibility to accommodate all of these divided lines.

Here is the point at which the philosophical study of racism itself becomes a kind of methodology, an approach for usefully thinking through other problems outside its immediate purview. If we gather up the insights from the literature on racism, we arrive at a positive launchpad for generating new ideas and resolving intractable problems beyond the scope of racism. This is the project of chapter 4 where I consider how racism is a springboard for re-framing the idea of responsibility. I continue the methodological theme of collapsing divisions, this time between responsibility models divided up into individual versus social camps.

Thus, beyond the more general approach this dissertation takes – philosophy as amelioration – another approach is borne out by the dissertation’s arc. Gathering up the themes and lessons of this dissertation reveals that the problems, phenomena, and analyses of racism together constitute a methodology in their own right – *racism as methodology*.¹⁷ If we expand the constellation of issues in the area of racism beyond the immediate, circumscribed purview of the racism literature itself, then we can ask: “How does racism apply to other

¹⁷ Of course, “racism as methodology” does not mean that philosophy should be done with any racist intent or motivation – that we should approach philosophy with a racist worldview. I am suggesting that racism can inform the ways we approach philosophical problems, and that it can constitute an approach for thinking about how those problems are resolvable (or not) and for generating new ideas with regard to existing views. So, racism should count as a philosophical methodology just as analytic philosophy, phenomenology, pragmatism, feminism, and so on, are all philosophical methodologies.

philosophical problems, and how does it constitute an approach towards resolving those problems, altering the ideas we work with in a general sense, and generating novel ideas?"

Thus, it is not just that the problem of responsibility is analogous to or parallels the problem of racism. I offer responsibility as an analogous example not only because it too suffers from similar narrow vs. wide, particular vs. global, dynamics, but also because the question of racism applies to responsibility in an important way – namely, with regard to the matter of privilege/power.¹⁸ Thus, racism as methodology – as an approach – just means that instead of using a moral lens to analyze racism, we reverse the order of analysis and use racism to analyze a moral idea – namely, that of responsibility. We frequently bring to bear moral concepts on racism in order to understand what it is. But why use what are mostly white, Euro-centric ideas of morality for our analyses? Instead, my suggestion is to flip the script, as it were, and engage those ideas of morality, generally, with what racism teaches us about the real world; hence, again, the emphasis on the ameliorative function of philosophy. This reversal, as it plays out in chapter 4, is meta-ethical in nature, asking how racism bears on our moral idea of responsibility, and what new things it might teach us. My point is just

¹⁸ But couldn't we, for instance, just switch out "race" for "gender" and come to the same conclusion about power and responsibility? We could, of course, but that is not the point I am making here. The idea is that racism is an approach to responsibility just as, for example, one can approach the notion of time phenomenologically. But that does not *ipso facto* mean that time is phenomenological, just as responsibility is not necessarily racist, racially unjust, or racialized. I therefore do not mean anything terribly sophisticated by *racism as methodology* – e.g., it does not mean rearticulating responsibility through the lens of critical race theory or critical philosophy of race.

this: why limit the impact of race-thinking within the realm of race and racism when it has much broader implications? If racism can teach us something about responsibility *independently* of race and racism, then that is all that is needed to constitute it as a philosophical approach.

So, instead of homing in on individual, shared, or collective agency – a dis-integrative approach – we ought to think of responsibility in terms of power. This provides not only a realistic lens by which persons and groups are held to account, but also reveals a plausible way to bypass the individual and the social where responsibility is concerned. Investigating the positionality of persons within groups along a continuum of power enables our tethering responsibility to the ground. I argue in chapter 4 that we get desirable consequences from doing so, leaving aside metaphysical complications and achieving a more nuanced view where we can more clearly identify who is more and who is less responsible for causing harmful events.¹⁹

3. Method

How does the foregoing inform the activity of philosophy? I think the problems drawn out so far are a poignant lesson for how to do philosophy. So instead of speaking explicitly in terms of what methodology is up to the task of resolving the tensions above *specific to race and racism*, I wish here to say

¹⁹ Harmful events would of course include racism, but only because my re-framing of responsibility is meant to generally capture any harmful event. Thus, the account I offer is not specific to responsibility *for racism sans phrase*. It is, rather, a general view of responsibility *inspired by* the problems and ideas of the philosophy of race and racism, namely a construal of privilege as a kind of power.

something much more general about how we should do philosophy. So, what I have to say is about method in general, full stop. Obviously, given that I am talking about approaching the activity of philosophy as such, it will apply to the problems of race and racism which fall under its umbrella.

So, what is my approach? What is the general position from which I think solutions to philosophical problems can be launched? I aim to develop below an approach in the spirit of Deweyan pragmatism, specifically in the spirit of reconstruction.²⁰ As a philosophical aim, reconstruction reassesses the approaches taken by other philosophical methodologies and attempts to offer a salve. It re-tools notions – for our purposes, *objectivity* and *generality* – so that they may be put to better use. But this is not done by sweeping away the whole edifice of philosophy. Nor does it mean entirely doing away with the insights of existing theories. Any reconstruction depends on what came before. The pragmatism I am after does not resist the incorporation of multiple philosophical methodologies to achieve practical results, where what is *practical* is what is best suited to the inquiry and problems at hand. Thus, while I will lean on John Dewey below, my intention is to make this approach my own, hence *in the spirit of* and not *by the letter*.

Philosophy as Amelioration

The point here is to emphasize the ameliorative activity of philosophy in the spirit of pragmatism – something I alluded to above – that things can be

²⁰ The chief influence here is Dewey (2004).

made better both in terms of philosophical theorizing *and* in terms of infusing hope into the real-world problems we face. I write “hope” not because I think there are any grounds for optimism but instead because I think there are only grounds for mitigation. But this does not entail that I am a pessimist. That philosophical theory and the concrete conditions of our lives can be made better does not mean that anything will be definitely *solved*, nor does it mean that anything will *never* be solved. Instead, the point is that amelioration is an ongoing activity; we do the best we can in addressing problems in philosophy and the world with what we already have to work with.²¹ Thus, the point of philosophy as amelioration is not knowing for the sake of knowing, achieving wisdom for the sake of wisdom, but *knowing for the sake of making things better* both in terms of theory and praxis. The idea here is that philosophy and the philosopher are social functionaries of humanity.

As we saw before, white philosophers must carefully climb uphill if we are to play a role in the philosophy of race and racism. One’s perspective can delimit what one can know, subjectively appreciate, and offer in a field so centrally about experience. It would seem that any attempt at speaking objectively is moot from the get-go. But my thinking is that we can maintain both objectivity *and* subjective perspectives – so long as a little methodological agility is allowed in the mix.

²¹ See Dewey (2004, 102-103) and Koopman (2006).

I have hinted already at the sense of “objectivity” that operates in this project, but so far only from within the context of race and racism. So, permit me to say more for the purpose of explicitly clarifying my non-standard usage of the term. What is needed is a notion of objectivity which does not calcify the (white normative-theoretical) status quo and does not pretend to be an Archimedean point. A “standard” sense of objectivity might mean some like a God’s-eye-view (or a view from nowhere) which is detached and decontextualized, where particular differences are ignored in the service of pure, unadulterated abstraction. The epistemic “advantage” here is that of a non-subjective perspective, that what we can know “holds” regardless of any one person’s idiosyncratic viewpoint.

But the sense of “objectivity” I maintain here and throughout this project turns the elements of the “standard” sense on their heads. To approach philosophy with an eye toward objectivity must *not involve ruling out disparate perspectives*. The point is that a plurality of perspectives is precisely what is *needed* for any inquiry to be objective. Information is gathered at the ground-level. Once the data is collected, the trick is to organize and order it so that we can speak in a way that anyone in principle could agree to. The reason agreement can be reached is that an agile notion of objectivity *preserves* difference. We can preserve differences not by fetishizing them, but by drawing out what seems to be common amongst them. After all, as Dewey (1922) notes: “...in certain fundamental respects the same predicaments of life recur from time to time with

only such changes as are due to change of social context..." (382-383). The point is to think through that recurrency, putting together the differences in a coherent, organized fashion so as to generate new ideas.

So, I take "objectivity" here to mean *general*, that various ideas and perspectives can be put together coherently, constituting a common ground from which anyone can analytically proceed and to which anyone in principle could agree. This does not mean objectivity is a Gods-eye-view, or that it is detached and decontextualized, or that it seeks to dismiss particular experiential differences. Rather, the sense I am using is *connected* to experience, a "bringing together" of a plurality of perspectives. It is therefore *attached, contextualized, a view from anywhere*, and appreciates particular differences. The epistemic advantage here is precisely that of collecting subjective perspectives, different theoretical points of view and the phenomena they capture, and that what we can know holds *because* of these particular viewpoints.

The agility of this approach to objectivity allows a negotiation of plurality and generality. What are we "putting together" if not a great number of differing perspectives, whether they be the perspectives of individuals or phenomena expressed by different theories? And what are the "recurring predicaments" Dewey references but kinds of "sequential bonds", identifiable threads of commonality that buy us abstraction and a general perspective on matters (Dewey 1958, 122-123)? We saw in the foregoing that there is a real tension between the particularism of the literature on race and racism and my stated

motivation to abstract away from the particulars so as to speak more generally. But if particularism is right, then must we do away with generality and abstraction altogether in philosophy? How is this to be negotiated?

In the same way that we need to collect differences to speak objectively, here again we need the particulars *in order to think generally*. The project of putting things together is only possible if space is made for alternative points of view, that is, if those alternative points of view are *retained*. So, pluralism/particularism do not preclude a general description, nor vice versa. (An endorsement of the general position *p* does not here mean *not-q*, *not-r*, *not-s...*, but instead means *not-not-q*, *not-not-r*, *not-not-s...*.) The proposed move is to incorporate into this approach both generality *and* plurality, where generality is a tool for thinking clearly only *after* the particulars are appreciated and gathered up.²² This does not mean *doing away* with the particulars, ignoring what we find in our context. We must start somewhere. But that does not entail generality is meaningless.

Nay, generality is a tool. It helps us grasp the particularities of our context (Dewey 2004, 86). The pluralism of my approach here is meant to recognize and respect difference without succumbing wholly to the side of the “many” of the classic “One and the many” problem. Instead, generality is *functional*, it brings “one” *out of* the many by abstraction (86-87). Abstraction, thus, is not a four-letter word. On the contrary, “[a]bstraction is liberation” (86), but not because it allows

²² I draw inspiration here from Dewey’s (2002) claim that “[t]o call a generalization a tool is not to say it is useless; the contrary is patently the case” (244).

access to some timeless, fixed truth. It is liberatory because it serves a *functional* purpose – it helps us to deal more effectively with concrete particulars.

So, the approach is agile, it doesn't preclude context, nor does it begin from a *universalist* position. It begins with particularity and builds toward something general, something common without pretending that the generality arrived at is set in stone forevermore. It helps us to find the best available approach for resolving the problems in front of us. We use the tools at our disposal where they help, so long as it is *for* something practical.

But "practical" here does not always and only mean "practically good for our social problems". (I'll say more about that briefly.) "Practical" for the pragmatist also, and no less importantly, means having consequences that are useful for inquiry (Dewey 1954, 330-331; Schwartz 2015, 36). Some theoretical tool is practical when it plays a problem-solving role both in thought *and* in the service of everyday life. The pragmatist need not be allergic to theory so long as it is practical in the sense described. Theory organizes and orders, it improves what we already know.

As noted above, we may rightly want to be wary of approaching problems with any specific metaphysical commitments in tow – the blurred and divided lines of amelioration and responsibility. The slipperiness of individuals, groups, and social structures present a point at which inquiry finds a limit. We get tied up trying to determine which of these should take pride of place. But instead of sitting with these tangles, restricting ourselves to either/or binary thinking, we

should put the metaphysics on hold and ask, “What tool should I use to move the inquiry forward?” Pick the tools that *work*.²³ This sense of “practical”, then, sets us on a productive, problem-solving path, leaving unproductive binaries behind and opening new ways to diagnose the issues.

Once we get our theoretical “ducks in a row” – i.e., once we have a general idea of what we are talking about – we can *then* speak to the practical ways the everyday affairs of the world can be improved. The philosophy-as-amelioration approach always involves a theory-practice cadence, and once we have cleared away the debris populating our thinking, our *praxis* will be well-informed.²⁴ That’s the point of philosophy – to be a public servant, a functionary, of humanity.

4. Previews

Chapter 2 – Racism, Affiliation, Amelioration: Why Non-Reductive Unification Matters

Chapter 2 faces the above tensions head-on, setting the stage for the rest of the dissertation. It argues that there is a role for the white philosopher, that she can be objective, and that metaphysically tangled amelioration strategies are resolvable. The central point it argues towards is that we ought to get descriptively clear about racism in order to make things better, because the way

²³ Of course, there are dangers in selecting any effective tool because doing so might commit us to one pathway towards a solution as opposed to others. That is, the tools we use can in many ways determine the outcome and proposed solution. But I show in chapter 2 that while this is an issue, it is resolvable.

²⁴ I use “praxis” here instead of “practice” to uphold the idea that what is practical is what has consequences for *theory* (i.e., for inquiry). I draw no strict boundary between theory and practice. Rather, the boundary is more appropriately drawn between *theoria* and *praxis* (in Aristotelian language), between contemplation and action/doing.

we describe can either limit or unshackle the way we prescribe. It gets there by making a case for how the diversity of views of racism in the literature can be unified without reducing them all to a single definition, that we can put things together while retaining difference.

It does so by investigating that upon which types of racism, and the phenomena they describe, conceptually depend; that is, what makes racism formally possible. It shows that this can be done via what I call the *affiliation diagnostic system*, a method of identifying our chosen and unchosen connections to groups. The diagnostic, it is argued, shows how any type of racism is possible. The importance of this method lies in its allowing anyone to think through racism from a position where unique first-person perspectives, which not all persons can possibly know, are not rejected. Ultimately, then, part of the value of “putting together” types of racism is theoretical clarity which in turn clarifies *praxis*. A better picture of the whole gets us a better picture of how to organize our strategies for ameliorating racism.

Chapter 3 – Overlooking: It Could Happen to You

Chapter 3 carries forth the “putting-it-together” theme of chapter 2, zooming in on more specific positions in the literature on unwitting racism. It shows that there is a bridge between two prevailing positions: *active ignorance* and *implicit bias*. That “bridge” is a phenomenon I call *overlooking*, which describes subjects who both care and know better and still undertake racist acts. This is all done through the perspective of the privileged white participant who,

although morally and epistemically competent, sees her racism nevertheless unfold in front of her. It argues that overlooking is irreducible to either active ignorance or implicit bias, the two models best equipped to capture the phenomenon. Though overlooking is not entirely active ignorance, it has a structural component to it, and though not entirely implicit bias, it is an individual phenomenon. While overlooking is irreducible to these models, I show that there are important continuities which reveal a new frontier for approaching racism, how social structures play out in individual psychology.

Beyond new directions for approaching racism, overlooking reveals undesirable moral and practical implications of implicit bias and active ignorance. I argue that these views enable moral evasion from personal responsibility and show that such evasion does not hold for overlookers. I end by highlighting some unrealistic practical implications of active ignorance and implicit bias, arguing that these views imply dangerous thinking that one is “done” with racism. I claim that overlooking shows no white person can stake a claim to such thinking.

Chapter 4 – A Continuum View of Responsibility

Chapter 4 draws out, very generally, the primary themes of chapter 3, showing that racism can function as a methodology for approaching other questions in philosophy, a positive launchpad from which new ideas and solutions can be generated. In chapter 3 we “looked through”, from a position of privilege, the perspective of the racist participant. We also saw in chapters 2 and

3 that racism is not so neatly reducible to individuals, groups, or social structures, and that there are good reasons for avoiding doing so. These two themes – privilege and irreducibility – motivate the arguments of chapter 4. Here we ask two questions: “What is privilege but a kind of power?” and “What use is there in thinking about responsibility through metaphysical lenses?”

The chapter argues that responsibility does not *just* have to do with individuals, groups, and structures and the ways they all interact. Instead, it has to do with how power operates as a smoke-clearing tool cutting through that interactive mess. Just as racism has been re-framed in the previous chapters as irreducible to either the individual or the social, we might do the same here vis-à-vis responsibility. Given that prevailing theories of responsibility generally focus on one or the other, responsibility too should be re-framed. Instead of busying ourselves with the metaphysical underpinnings of responsibility, by focusing on power we arrive at a much more nuanced view – that responsibility lies on a continuum of power. In virtue of their position within a group, some are more responsible than others.

Conclusion – Charting Out New Paths: Two Sites for Future Research

The dissertation’s conclusion drafts some ideas for future research borne out by some of the dissertation’s major themes. There, I carry forth the themes of chapter 4 (the responsibility-power continuum) and re-direct attention *back* to privilege. I unpack the idea that responsibility *itself* is a privileged affair and sketch a general position according to which responsibility is asymmetrical, that

it is a disproportionately available privilege. The general view is supported by the socio-economic reality that privilege is involved in the capacity to *take up* responsibility *at all* for oneself, one's actions, and for others. Doing good, that is, is available only to those who have the *means* to do good.

I then consider, again, the role of the white philosopher in the philosophy of race and racism and the aggressor/oppressor perspective. I ask whether there is more to be learned about white psychology from the white perspective. What I propose, first, is to approach white psychology from a non-moralistic starting point, and second, to approach microaggressions from the perspective of the white aggressor. The point is that if we don't look through the perspective of the aggressor, we will miss out on a realism about her psychology, and what we get is a subtler, greater range of the phenomenon. In the style of overlooking, it is a diagnostic for those who don't want to engage in microaggressions from the perspective of the aggressor.

Chapter 2

Racism, Affiliation, Amelioration: Why Non-Reductive Unification Matters

Introduction

Racism, at every turn, resists order: it can be caused unwittingly by one's psychology, live and thrive in social structures, depend on different epistemic points of view, be contingent on different histories and locations, and may be describable theoretically or moralistically. But which point of view is the right one? Is there a right one? If there isn't, it seems we are left with a mess. With so many moving targets where it comes to racist phenomena, when we want to hit those moving targets, we end up with a fragmented multitude of theories of racism. There just seems to be no *general*, basic idea of what we're talking about. And because of the fragmented views in the literature, each view comes up with separate ameliorative strategies; each, that is, offers different ways to lessen racism. But if we cannot get straight on how to get a picture of the whole, how do we mitigate the various instantiations of racism, or indeed racism as such? We should want at least some bare minimum commonality between them so that we can effectively approach how it all might be made better. But, what's the alternative to a multiplicity of theories of racism? Perhaps we reduce racism to some one thing, even if that misses out on the subtleties of real racist phenomena. There is a dilemma to face up to: either accept a multitude of fragmented theories, with no coherent concept or ameliorative strategy, or just

reduce it all to some one thing, ignoring contextual differences and ameliorative strategies.

This chapter asks how we might approach unifying theories of racism in general and attempts to dissolve the dilemma by showing that there is a way to be objective in the philosophy of race and racism, a field resistant to objectivity, *while* respecting the differences of varieties of racism and without being reductive. But *can* a white philosopher (i.e., this author) be objective in this field?¹ What is the role of white philosophers in this conversation and what are their limits? It is easy for whites to not know or talk about aspects of racism because there are no bad consequences for them. Yet when I stake a claim to objectivity, I'm saying that there is a way to think about things that anyone could in principle agree to regardless of their perspective. But what becomes of first-person perspectives if we endeavor towards an objective point of view? Are they to be ignored for the sake of descriptive simplicity?

One way to objectively analyze racism is to reduce it all to a single definition: racism *is* *x*. But if a philosopher of color understands the consequences of some racist phenomenon as stronger or deeper than whatever is captured by *x*, then on an objective approach *qua* reduction, we may be required

¹ Recall from chapter 1 that the sense of "objective" is not the "traditional" sense of the word—that is, it does not mean anything like a God's-eye-view. As one will see in a moment, this does not mean I am asking anyone to adopt a homogenized white perspective. That is, I am not asking anyone to "see it like whites see it". Rather, what I am after is a perspective open to anyone at all. Moreover, I am *not* saying that because I lack the experience of racism, all that I can do is assume an objective stance. What I have to say here and below should *not* be confused as conflating objectivity with white normativity or "white" analysis.

to deny that first-person perspective. But I'm not attempting to sanitize the scene. I will try to be objective in a different way that puts the pieces of racism together as a whole *without* reduction and *with* retention of first-person perspectives.

To fill these tall orders, I'm going to take a meta-theoretical second-order approach: an analysis of analyses. What I mean is that I will *look at*—i.e., analyze—analyses of racism and try to make sense of it all. That is, at a second order, I am offering a way to think *about* how first-order theories explain racism.² I take up this second-order perspective to identify, objectively, what order there is in theories of racism and the phenomena they capture. What being objective means here has to do with *approach*, a *way of thinking* conceptually about racism at a meta-theoretical level. I do so by invoking the tool of transcendental analysis.³ All I will mean by this is *thinking about what makes racism possible*, that upon which all the different types of racisms (and their phenomena) *conceptually depend*—i.e., how it is possible that there is racism *at all*. I do not mean this in any metaphysical sense, that theories of racism and their phenomena *are* such-and-such. Nor do I mean to suggest that this is a program for “correctly” or

² To be clear, I am not suggesting that a second-order approach is synonymous with whiteness or that it is an unraced “ideal” to which we should aspire. Nor am I saying that *all* whites can do is a second-order approach. The approach is not closed to anyone at all. The very idea that I would mean that this meta-theoretical, second-order approach is *just* for white philosophers would imply racist connotations—namely, that nonwhites cannot do meta-theory. Obviously, I mean the exact opposite. The point is just that we need a second-order approach to get a sense of racism as a whole for ameliorative purposes.

³ This may seem odd given my pragmatic approach explained in chapter 1. However, it is not inconsistent to take a transcendental approach here. Taking philosophy as amelioration *as inspired by* pragmatism does not preclude using other methodologies at our disposal. We use the tools that work for the problem at hand, and I argue below that this is the best available tool for the analysis I conduct in this chapter.

definitively answering how to describe racism. So, I will not identify a novel phenomenon, advance a new first-order theory, or explain what racism *is* in some way that connects all the views together.

What I *will* do is approach differently how we think about theories of racism, a new way to approach it to which anyone in principle could agree. I will show that this can be done by what I call the *affiliation diagnostic system*, a method of identifying individuals' chosen (or unchosen) connection to some group(s) which shows how any variety of racism is possible. The import of this diagnostic method lies in its putting together racisms in a coherent non-fragmented, non-reductive way. It retains the varieties' differences and respects an epistemic perspective that I don't have, all while not losing sight of the unique value of the respective theories that may be lost via reduction. The reason this approach is important is that it will enable anyone to think through the whole of racism from a position where one need not reject authors' unique first-person perspectives which not all persons can possibly completely know.

But beyond non-reductive unification and its benefits, why should it matter whether there is any order to racisms, and why should we want objectivity at all, looking at things at a second order? These things matter because it helps us to make racism better *as a whole*. Instead of the way we *describe* limiting the way we *prescribe*, we can unburden ourselves and use any amelioration tool at our disposal. Beginning from a sound generalist framework allows us to get our ducks in a row in terms of real practical ways to address the

situation so that we can effectively leap into praxis. Ultimately, then, this is a chapter about method. It is about how to make our understanding of racisms *theoretically* clearer, in general, and show that clarity of the whole helps clarify *praxis vis-à-vis* strategies of amelioration.

1. Putting Together the Picture

This section argues for a dissolution to the dilemma. It shows that we need not hang our hats on either the side of fragmentation (horn 1), by which I mean a multiplicity of independently valid theories and phenomena, or reduction (horn 2) by which I mean an approach which boils it all down to one thing. I will call these horns the “many” and the “one”, respectively. I argue for non-reductively unifying varieties of racism and the phenomena they capture by putting them together in a coherent way. I take cues from Martha Nussbaum’s (1988) method of identifying a ground underneath culturally different actions, something basic on which they all depend. This launches my own transcendental approach, a formal way of thinking at a second order (clarified in section 1.3) which focuses on what makes racism possible *at all*. The reason this approach matters is that it dissolves the dilemma, unifying racisms without focusing solely on difference or reduction. Differences are retained. Theories are non-reductively put together. And the approach results in a valuable tool, the *affiliation diagnostic system*, which helps us to understand varieties of racism as dependent on different ways we are connected to groups. Ultimately, this tool is valuable

because it allows us to talk objectively about racism without denying first-person perspectives and without oversimplifying the conversation.

1.1. Horn 1: Selections of the “Many”

Let us begin by tackling the first horn of our dilemma – accepting a multitude of fragmented, satellite theories of racism. This sub-section is our base camp from which our expedition towards putting together racisms begins.

Though I will draw out some consequences critical of the fragmented landscape, I do not mean to deny the positions’ unique value.⁴ We should appreciate the richness of each position’s identification and explanation of its own particular form of racist phenomenon, because identifying the various ways racism manifests provides inroads towards identifying what is variously morally problematic about it. Nevertheless, what do these “satellite” racisms all have to do with each other?⁵ Shouldn’t we want to know how or whether they all work together? The question to consider in this sub-section, and one to which I explicitly turn in section 1.3, is this: do they have anything *at all* in common?

⁴ Allow me to dispel a possible confusion with regard to the dilemma’s framing and what is stated here. I do not mean to suggest that the consequences of each horn – the “one” and the “many” – are *equally* negative. I present over the course of sections 1.1 and 1.2 a picture according to which the “many” has less severely negative consequences than the “one” (of which I am more critical). But this should not be taken to imply that the “many” does not have issues of its own.

⁵ To be clear, I do not presuppose, nor will I say, that racism is just one thing, nor that it must be defined with necessary and sufficient conditions. It could be constituted by family resemblances or by multiple models. The fact that there are many racisms is not a problem, and I have far more sympathy for this than I do for a conceptually defined RACISM (but see note 6). Because my goal in section 1.3 is to integrate racisms – that the theories and phenomena should be *retained* – I have no reason to reject multiple views. Thus, unlike those who might want to conceptually define racism, in which case they would have to exclude certain models of the target phenomena of racism that would fit their intension of racism, I neither must nor will do that.

Now, as I cannot possibly go over all varieties of racism in the literature, I will focus on some views I take to be representative of fundamental themes in the literature. I take up Garcia's (1996, 1999) volitional account of racism (VAR), active ignorance (AI), and implicit bias (IB). The reason I pick these views is that they cover a wide swath where racism can be construed as ill-intentioned motivations (VAR), the result of willful ignorance which maintains structural racism (AI), and nonconscious action (IB). The thing to notice is that each position points out something uniquely different about racism, something that is linked to each position's basic assumptions and commitments.

Let us begin with VAR where, according to Garcia (1999), racism is always a matter of "racial dis-regard or even ill-will..., the core of the phenomenon" (13). Though individualistically-centered, Garcia thinks that VAR covers a wide range, from individuals to institutional practices (13). Importantly, however, whether it be a racist action or institutional practice, racism always starts at the "hearts" of individuals (Garcia 1996). Within its purview, VAR picks out a genuinely racist phenomenon. We need look no further than those who harbor explicit racial hatred towards someone or some group to see that VAR uniquely captures such persons – they are not unintentionally or unwittingly biased. And we would not want to do without VAR, because it helps us understand the egregiousness of overt individual racism.

Does this capture all racist phenomena? VAR is problematized by AI, a view which adds more subtly to a description of how racism is generated and

maintained. While we do not need to disagree with Garcia's VAR, there is much more complexity to explore beyond disregard and ill-will, namely those more covert aspects of racism. AI identifies a "structured" willful white ignorance as a site from which racism is produced and maintained. Whether the discussion centers on resistance to acknowledging racism or knowing anything about it (Mills 2007; Bailey 2007; Medina 2013), or a deficient *motivation* to acknowledge or know that whiteness maintains a racist status quo (Mills 1997; Sullivan 2006; Alcoff 2007; DiAngelo 2011, 2018) what is common amongst these views is a point about *deliberate* ignorance of racial injustices and inequalities, with the additional goal of furthering white supremacy and maintaining power. Like VAR, AI uniquely picks out something genuinely real about racism, namely how it is sustained at both individual and structural levels. Though AI emerges from whites' epistemic resistance, its "value" is its broader social impact, an attempt (conscious or not) to uphold white privilege and supremacy, to "hang on" as it were, to the benefits whiteness affords white people. We should want to retain the AI position as it betters our understanding of how white denial of racial injustices buttresses and preserves racist social systems.

But there are aspects of racism which are not the result of active (or non-active) ignorance perpetrated even by those who know better and care a great deal about *not* doing racist things. Could knowledge and care stand as stalwart barriers against racism?⁶ Not according to IB. Like VAR, we may focus

⁶ I write both "knowledge" and "care" because one might know a great deal about how and why she should avoid race-based harms but not be motivated to avoid those harms – e.g.,

individually, but instead of searching for racial disregard or ill-will, the point here is to unpack someone's implicit biases. Even the most well-intentioned people fall prey to automatic associations made within their psychology, relying non-consciously on unfair racist stereotypes to guide their behavior and actions (Kelly and Roedder 2008; Devine et al. 2012; Holroyd 2012; Hardin and Banaji 2013; Banaji and Greenwald 2016; Amodio and Swencionis 2018). IB presents a picture where unwitting racism is sometimes right around the corner. What IB uniquely captures is that racism can result from psychological processing, that it can catch us by surprise. Retaining this position is important because without it we would lose a distinctly empirical understanding of unwitting racism, as well as an integration of social and individualistic scope.

Now, given that I think the foregoing should be retained, I have no reasons to reject any of these *positions*.⁷ But there are some worrying methodological consequences if we accept that making sense of racism amounts to offering a multitude of fragmented theories. It is the *approach* which should trouble us. For instance, where we see some phenomenon *p* that may or may not

knowledgeable sociopaths. The point is that there needs to be *some* motivation to avoid racism, even where one is knowledgeable.

⁷ One might wonder at this juncture whether active ignorance or implicit bias are accounts of racism *per se*. I think there are at least two ways to approach this question. First, challenging that AI and IB are not accounts of racism *per se* may be a politically or morally motivated move to limit "racism" to an intentionally narrow range of phenomena that allows for evading being labeled a "racist". Second, and I say more about this in note 8, theorists of AI and IB, to my knowledge, are not positing that "Racism" with a capital "R" *just is* active ignorance or implicit bias—i.e., they are not *defining* "racism". They can be understood as *models* of racism. That is, they are meant to explain or "pick out" certain kinds of racist phenomena or phenomena with racist connotations and implications.

be racist, we tend to go to our preferred method of description to determine whether p is racist. Even though these positions do not take a reductive *approach* in general (i.e., reduction is not their *aim*), when viewing p through the lens of their position, there is a kind of practical reduction afoot.⁸ What I mean is that for the purpose of explaining p , a theorist of AI might see p as an instance of AI whereas a theorist of VAR might see (the very same) p as an instance of VAR.

The point is that theorists attach their positions to understanding particular cases. Take for example the “All Lives Matter” response to the Black Lives Matter movement. This response is arguably best explained as an instance of AI, a way of deliberately ignoring the Black Lives Matter message so as to deny its import and own up to the role that whites play in making it such that black lives do *not* matter. But if one holds sacredly to VAR, she may see “All Lives Matter” as an instance of ill-will, hate, or disregard directed towards blacks because they are blacks.⁹ The point is that if we are so attached to our preferred model, then we run the risk of misidentifying phenomena.

Though this fragmented landscape is not a hyper-competitive one – i.e., it is not an “I’m right, you’re wrong” environment – there is a narrowness of approach these positions take. That is, none I think would reject the others as flat-out wrong – e.g., being an IB theorist would not require denying AI or VAR

⁸ Though I am saying *defenders of position X see p as X even if it seems p is actually captured by position Y*, this does not mean X-ists say *all* racisms are X – e.g., to my knowledge, no theorist of AI says all racism is born of, or just is, AI.

⁹ We might well add “contempt” to this list. For Garcia (1996, 6-7; 1999, 13-18) “contempt” is included in the notion of disregard.

wholesale. But even if there is no “my way or the highway” attitude, the practical reality is that this attitude is where we end up when we accept an overall method of *specialization* without any generalist understanding. We end up there because, unique phenomena require unique theoretical insights and explanations. This is the stuff of a specialized approach, concentrating hard on some particular aspect of the racist whole. It is inevitable, then, that we have a fragmented landscape; a lot of unique phenomena means a lot of unique theories. But the result is a methodological narrowness, a level of specialization that is required for making sense of the complexity of racism. And so, beyond the issue of possible misidentifications, I want to raise a further worry relating to methodological overspecialization. Experts offer views so narrow that a generalist view falls outside the scope of investigation. Who are these views for? It seems very likely that they are for other academics, resisting as a matter of method what would otherwise be helpful for the public’s understanding.

My point is that overspecialization is not making matters clearer for the people who are (or should be) our targets – the public – and it becomes difficult for the public at large to trust what experts say. The public can deny the nuances of structural racism wholesale because one likely has to be an “academic” to understand them. They can reject IB because it assumes special knowledge of psychology. The point is that specialized discourse alienates the public because it requires specialized academic knowledge. The role of the pragmatist (or of philosophy as amelioration) should be to overcome specialization and invent a

different way of approaching racism that provides a function for the public. We should help clarify, not confuse, because racism is a problem of the public for whom we should offer clear tools for understanding and combatting racism. What could be more ameliorative than something the public could understand and get on board with? If we grant that the role of the philosopher is arguably, at least in part, to offer theoretical clarity for the sake of the public (i.e., to be a functionary of humanity), then we should tone down the specialized discourse and offer broader brushstrokes.

1.2. Horn 2: Reduction, or the "One"

If we want to avoid the consequences of horn 1, and paint with a big brush, then how might we think through a more general framework that could put these views together? One option, and the second horn of our dilemma, is to reduce racism to some single explanatory feature. So, if we want to put it all together, then perhaps we could reduce the different varieties of racism. The reductive approach is appealing because it simplifies the fragmented landscape in such a way as to avoid butting our heads against a brick wall, especially where it comes to capturing conflicting domains (e.g., *individual* and *structural* racisms). Though I have said that my aim is to be objective, I will argue that achieving this aim via a reductive approach is unattractive because it obscures more than it coheres, generates an interpretive double-bind, and likely denies non-white first-person perspectives.

Whether it is a “desire to dominate” (Schmid 1996), ideology (Shelby 2002), “inferiorization” or “antipathy” (Blum 2002a, 8; 2002b, 210), or “disrespect” (Glasgow 2009), all these positions share or imply a common approach, a sort of reduction or at least a reduction in spirit. By that I mean where we want a satisfying notion of “racism”, we simplify the various complexities to find a common core – where such simplification has the effect that racism, metaphysically, *is x*, or conceptually, is understood *as x*.¹⁰ Being reductive “at least in spirit” just means that there may be something missing that does not fall under the racism umbrella, but that there are good reasons to think in reductive ways about racism – e.g., for simplicity’s sake, greater explanatory power, or a satisfying extension of the word *racism*. The distinction between *metaphysical* (racism *is x*) and *conceptual* (racism *as x*) simplification matters because not all reduction is monistic (as in “racism *is this one thing, x*”). Some theories offer a conceptual strategy to make sense of racism by analogizing it with something else (as in “racism *is like x*” and so we can understand racism “*as x*”). The distinction is important because we should not confuse all reduction as metaphysical simplification.

¹⁰ I stated in this chapter’s introduction that I will offer a way of thinking *conceptually* about racism at a second order. To dispel any confusion about my claim there and my labeling here of conceptual reduction (i.e., racism *as x*), allow me to clarify by drawing a further distinction between *conceptual in the first order* and *conceptual in the second order*. I take those who offer descriptions of racism *as x, y, or z* as doing first order conceptual work to capture racist phenomena. But I am after, at a second order, that on which all the different *x*’s, *y*’s, and *z*’s conceptually depend. The difference between the two will be spelled out in more detail shortly.

What I am going to do here is explicitly address Glasgow (2009), because he offers a unification view for racism (64). So, we will look at Glasgow's "disrespect analysis" of racism (DA), but not to determine whether he has the right *content* of racism (that it is fundamentally disrespectful), but rather because he has a representative reductive *approach*. Since Glasgow is explicitly going for a unification of theories of racism and the phenomena they capture, I am singling him out as representative, because I too am aiming for unification, albeit non-reductively. The aim here is to show the shortcomings of the *reductive approach in general*, not just the specifics of Glasgow's view.

Seeking to explain ordinary usage of "racism" (92), Glasgow's DA may provide the key to capturing all racisms. On the DA, " φ is racist if and only if φ is disrespectful towards members of racialized group R as R_s " (81).¹¹ Like Garcia's VAR, Glasgow takes his view to be powerful enough to explain the basic character of any instance of racism whether individual or social/institutional (82, 91-92). DA, though, has an upper hand over VAR because it does not depend solely on the hearts and minds of individuals.¹² Indeed, disrespect is broad enough to describe inter-personal and institutional aspects of racism and racist societies (82-84). For instance:

¹¹ He uses "racialized groups" instead of "races" because he wants to remain agnostic about the question of whether race is real (81). Racialized groups for Glasgow are "groups of people who have been identified and treated as if they were members of the same race" (81).

¹² Hence my focusing on the DA as an example of the reductive approach as opposed to VAR. The reason that VAR is categorized under the "many" instead of the "one" is that it offers a *specific kind* of racism. The VAR does not aim to understand all varieties of racism as "in the heart"; rather, it implies that all other varieties simply do not get racism right, a denial of racisms that are not in the heart. The difference, then, is that the DA unifies by reduction, that all varieties of racism are *unified* under the DA.

Antimiscegenation laws failed to respect rights of intimacy and love and the people whose rights were at stake. The political, legal, and military institutions in the United States that collectively enabled the near extermination of multiple indigenous American peoples disrespected them on several fronts, not least as members of sovereign nations and as persons with a moral standing that entails rights of life and security... (83-84).

At whatever level, then, disrespecting someone or some group on account of race could be a descriptive root to any racism whatsoever.

But while a reductive analysis offers simplicity and explanatory breadth, there are reasons to be wary of the approach for the purpose of putting racisms together. It creates tunnel vision where views like the DA can obscure our sight. The approach misses out on the sophistication and subtlety of other views because some aspects of racism fall through the cracks in reduction, aspects which are not neatly packageable by disrespect. Take, for instance, “racism as self-love” (Silva 2019) which describes whites as wanting to maintain, as a matter of self-preservation, the privileges they have at the expense of non-whites.¹³ It is difficult to see *where* the disrespect is in this, because whites are not racist on this view *out of* disrespect, but instead, as a means for protecting their social status.¹⁴ Neither is it just a matter of unintentionally “disrespecting” non-whites, because the “self-love” *aim* of whites as protecting their status is not necessary intentional

¹³ Racism as self-love is not the only instance of “racism as *x*” that is not neatly reducible to the DA. Further instances might include, for example, anti-black racism as bad faith (Gordon 1995), racism as technology (Russel 2018), and racism as “structural domination” (Omi and Winant 1994).

¹⁴ This challenge does not emerge solely from Silva’s account. This sort of reductive approach will not help elsewhere. It is unlikely that disrespect can capture the nuance of automaticity in IB. AI seems to be a case of willful disregard with ulterior motives of maintaining white supremacy and power rather than disrespect. VAR includes hate which seems, *prima facie*, stronger than disrespect.

either. It is, whether intentional or not, that whites exhibit racist behaviors with the *aim* of protecting themselves rather than the aim of harming or “disrespecting” others.¹⁵ And so, it would be misguided to ask whether self-love racism just amounts to disrespect.

This difficulty relates to a further worry that the DA approach will lead to awkward interpretations of or the discounting of views like self-love racism, attempting to put them in the disrespect “box”. Supposing we put our disrespect-tinted glasses on, where different varieties of racism come into view, we may end up in an interpretative double-bind. Either we awkwardly interpret those varieties by gymnastically rationalizing them as an instance of disrespect, or we discount those varieties as not “genuinely” racist against good reasons that they *are* genuinely aspects of the racist whole. And if it really is all disrespect, then how do we mitigate racisms that do not amount to disrespect? If the only tool you have is a hammer, everything looks like a nail.

Though I have said that we should want an objective overview of racisms, the only way that might help us achieve this — this reductive methodology — seems unattractive. Being objective via that method would seem to require us to deny *subjectivity*, that is, the epistemic privilege that imbues many notions of racism. And striving for something universal may require our ignoring the nuance of particularity where racist phenomena and the theories that capture

¹⁵ The same could be said of institutions. That is, it is not just that they are not explicitly, intentionally doing racist things, but rather that their motivation for doing so (which might be self-love) is different than disrespect supposes.

them are concerned. Should we kindly tell George Yancy (2008) that his experience of the “elevator effect” really boils down to disrespect? The point is that reduction as a *methodology* is tone-deaf, it runs the risk of ignoring the value of perspectival knowledge that theorists of color afford. If I, a white philosopher of race, want to offer an objective overview of racisms, then I will need to avoid the consequences of a reductive approach.¹⁶

1.3. *Dissolution: Method and the Affiliation Diagnostic System*

My aim in this section is second-order clarification of what theories of racism and the phenomena they capture have in common. To wit, I will offer a new way of thinking, just a framework, that can avoid the methodological issues of the “one” and the “many” and yet retain their methodological virtues. I will unify without reducing and without denying the fruits of overspecialized narrowness. My approach is a formal, generalist second-order meta-theory that does not rule out first-order analyses of racism.¹⁷

¹⁶ One might worry, à la Blum (2002a, 1-32), about watering down a theory of racism to the point where it is ineffective. So, let me be clear: any definition of racism I entertain is heuristic – i.e., for clarity’s sake for the present inquiry. They work as operating definitions for the purpose of developing a general second-order view of the first-order landscape. To reiterate, I am not offering a definition of racism, but an overview. An *overview* supervenes on what is already there, and it is not reductive. There is a trade-off between accuracy and explanatory power. A reductive approach sacrifices the former for the latter. So, the worry *ought to be* about missing certain phenomena in sacrificing detail and accuracy for explanatory power. My thinking is that there is nothing wrong in being accurate to the many different ways racism manifests, and that there are no good reasons to prefer reduction over accuracy. So, what I am after below is a formal “core” of racism which allows us to understand why it manifests in various different ways.

¹⁷ Such generality can in fact help in coming to appreciate a *novel* first-order articulation of racism even if such an articulation proves to be an exception to the “rule”. Indeed, whether the generality fails may also be a worthwhile investigation. That is, if a new first-order articulation does not fit with the meta-theory I offer, that is not a death knell to the meta-theory, because then we have the opportunity to look at the differences which change the whole (and not the other way around).

By *first-order analysis*, I mean an explanation of some phenomenon *p* as this-or-that, or an explanation that *p* is this-or-that, providing the content of *p* for the purpose of “picking out” *p* from amongst the crowd.¹⁸ And by my own *second-order analysis*, I mean analysis of first-order analyses. This involves thinking *about* how first-order analyses explain racism – their nature, their goals, and their methodologies. So, my approach is *meta-theoretical* as opposed to theoretical; that is, I am not offering a first-order theory explaining some racist phenomenon, but instead, a second-order (meta-)theory of theories that purport to explain racist phenomena.

Now, nearly all varieties of racism seem to involve some notion of race, harm, and the way people are grouped and interact. So, there ought to be something common running through the varieties based on these first-order contents, something to which we could reduce them all. But we saw the problems with reduction. So, my suggestion is to think about the varieties of racism *formally*. The formality of my approach takes a cue from Nussbaum’s (1988) method, which is to find the lowest common denominator of human action, that on which all varieties of action depend.¹⁹ Her process is to compare varieties of actions – seeing whether there is a common overlap amongst them –

¹⁸ For instance, an explanation of racism (*p*) as “self-love” or that it (*p*) is disrespect.

¹⁹ Nussbaum’s question is the particularity versus universality of ethics. While this has some import for racism, I consider solely her *method* and not that issue of content. As Nussbaum’s method is to allow for culturally particular flexibilities while finding a universal core of experience, I distance myself from that specific question, whether racism is universal at all or merely historically and culturally relative. The point here is about how varieties of racism (theories and phenomena) all depend on a basic foundation like Nussbaum’s “spheres of experience”.

for the purpose of finding what they have in common. It is by comparison that we find a lowest common denominator, much like if we want to know whether there is some form to triangles. We never “see” a perfect, pure TRIANGLE, but by comparison of triangles, we come to know that there is some form underlying them all. The result is some form that makes any action possible *at all*, a “sphere” of experience in which we must act. The point is that identifying commonality reveals what could formally *account* for that commonality, that we can allow differences without losing sight of what is *basic* to them all. Though we never “see” this lowest common denominator in any one action in any one culture, the overlap between the local variations is found by comparison.²⁰ So, by *form* I just mean that on which all first-order content depends. I understand *dependency* in transcendental terms where varieties of phenomena *p, q, r...* and the first-order theories which capture them are all made possible by some basic formal ground. A formal transcendental analysis of theories of racism and the phenomena they capture will help us think about how it is possible that there is racism *at all*.

How, then, will this resolve the problems with fragmentation and reduction? I argue that there is a unifying form on which all first-order content depends, thus not ruling out any first-order content, all while putting the pieces together – unification without reduction. I will do something like Nussbaum but

²⁰ Now, we could stop at comparison vis-à-vis racisms. But if we stop there, then we are reducing, because identifying a common overlap implies a first-order “core” or extension of “racism”. That is not my aim, and indeed, I’m going beyond that at a second-order to show how to unify racisms without first-order reduction or extension, all while not presuming to know better than those with epistemic privilege.

with racism as my target, offering a formal ground on which racisms all depend. Just as there is, objectively, *some* way we must act (a sphere), there is, objectively, something basic on which racisms (theories and phenomena) depend, *affiliation*, a common and necessary feature of human experience. All I will do is chart a middle path, a meta-theoretical framework that puts the first-order pieces of racism together in an objective, coherent way so that no one is just wrong – retaining epistemic privilege – but at the same time not reducing. My framework will not say what racism *is*, at a first-order, but offer a basis for uniting racisms at a second-order. Thus, none of this will help us to “see” the essence of racism, and it is likely that there is no “pure” reductive essence of racism – rather, the basic foundation is generated by comparing all the different first-order theories of racism, just as Nussbaum generates a basic foundation (the spheres) by comparing different kinds of action.

1.3.1. *The Affiliation Diagnostic System*

Given that racism *basically* has to do with individuals and groups, maybe there is something to just looking at the relationship of an individual to a group. And when we just look at this more basic frame, we see that if it were not for the relationship of an individual to a group, racism would not be possible.²¹ This offers a more basic way to think about racisms that can unify all the varieties. So, we are headed in a new direction: showing that the varieties of racism can be unified by explaining what makes them all possible, namely subject-group

²¹ While it is an interesting question whether someone could be racist to herself, I am agnostic about this and thus leave the question aside.

connection, *affiliation*. We will see that the different ways individuals affiliate map onto the different varieties of racism. Now, it is possible to misunderstand me as making a series of metaphysical points below, but recall that I am talking about formal transcendental conditions of possibility of racism. This is just a meta-theoretical diagnostic system that gives us purchase on how all the positions in the literature on racism fit together as whole.

Before seeing what the affiliation diagnostic system does, some technical terms on which I rely need straightening out. First, an *affiliation* is that which makes possible a connection between a subject and a group, and/or an attachment someone has to the set of beliefs constitutive of that group. Though the content of one's affiliations to different groups likely differs, *that one has* affiliations is true of any subject; that is, we all belong to groups regardless of the specifics.²² So, it is at least *a priori* plausible that the "attachment" we have to groups is formally basic in the sense that it is a part of the experience of any one of us, whether recognized or not.²³ Now, affiliation as such can be spelled out in

²² Cases like feral children notwithstanding, I am assuming that all subjects are connected to groups, and that there must be groups to which subjects can affiliate. I am assuming, also, that the sorts of things that in part make us human—e.g., a sense of self, language, sets of values, beliefs, aims, and so on—likely manifest from group membership. While feral children are obviously human, they miss out on some of the basics of human experience. I am just offering a technical definition that is likely safe to assume.

²³ But what makes affiliation specifically a condition of possibility of *racism* when it is just as plausibly a condition of possibility of friendship or religion, neither of which have much (if anything) to do with racism? There is nothing unique (in terms of first-order content) about affiliation that makes it a condition of possibility of racism beyond its being presupposed in our thinking about race relations. While it is likely true that certain affiliations come with disproportionate access to material recourses and socio-political capital along racial lines, I am not making that claim. Doing so would amount to first-order conceptual reduction: that racism should be understood *as* affiliation.

at least two distinct ways: groups of which we *do not* and *do* choose to be members. I call the former *involuntary* and the latter *voluntary* affiliations.

By *involuntary affiliation*, I mean affiliations which are accidents of birth, a passive connection a subject has to some unchosen group. The “accidental” nature of involuntary affiliation can be exemplified in any number of ways, say, by one’s connection to family, sex, disability, race, to those who are 5 feet 5 inches tall, and so on. Though these affiliations are unchosen, the connection a subject has to these groups likely influences her psychology, how she thinks of herself and others given the milieu of sociocultural features associated with those groups into which she was born. Whether wanted or not, involuntary affiliations are a probable source of our unnoticed and unreflective beliefs and values, of which we might adopt some or all. Yet, if all affiliations were involuntary, then some groups would not exist – bowling teams, friend groups, political parties, and so on.

Affiliations that are not involuntary are *voluntary affiliations*, by which I mean a subject’s chosen group attachments. Common to any voluntary affiliation is *activity*. The activity of voluntarily affiliating with some group allows us to “go beyond” our involuntary affiliations. For example, when going beyond what is associated with an involuntary affiliation to *whiteness* (and its possible racism and benefits incurred by an accident of birth), that might mean being affiliated with something like “wokeness” (acknowledging those benefits and possible racism to lessen them). I take “wokeness” to be a shared set of beliefs constitutive

of the *woke* group, and so in a case like this, the subject's affiliation to "wokeness" is *ipso facto* an affiliation to the *woke* group. But this is possible only if subjects actively develop affiliations outside their in-groups as a matter of *conscious choice* and does not mean simply "being around" subjects of an out-group. Those who do so are likely sufficiently motivated because they care about others and/or about who they want to be (e.g., whites who want to be antiracists).

If we grant that affiliation is basic to human experience, then against that backdrop we can see that theories of racism and the phenomena they capture have affiliation as a lowest common denominator. No variety of racism is possible without it. We can see how fragmented theories of racism are not as disparate as they may at first blush seem, and we do not need to be reductive to clean up the picture. Instead, we can clean things up by diagnosing the sorts of affiliation that make racisms possible, thus showing that we can unify the whole in a coherent way. As in section 1.1, I cannot possibly be exhaustive, encyclopedically reviewing all the varieties of racism in the literature. But if my view holds water, then any first-order position on racism is mappable by affiliation.

The Consistent Involuntary & Voluntary Case

Let us begin with a case where a subject's involuntary and voluntary affiliations are the same. Where there is no difference between these affiliations, we can clearly think through what makes a notion of racism such as Garcia's VAR possible. We can understand VAR as mapping onto subjects who both

absorb the forces of privilege and power given an involuntary affiliation to whiteness, and who do not disaffirm at a higher-order the beliefs and values that result from those forces.²⁴ The reason we talked about involuntary affiliation was to capture the way sociocultural influences of certain group memberships have on members' psychology. But in the case of VAR, does it matter that structural privileges and power are the cause of subjects' racism? Regardless of the source, it is clear that there just is no difference between these racists' nonconscious and conscious beliefs, intentional or unintentional forms of racisms, because they outright endorse racism.

The point is that their voluntary and involuntary affiliations are consistent. There is no conflict here and no difference. In a word, there just are people who are actively, pridefully racist. They do not go beyond the influences of their involuntary affiliations, instead embracing them. By "go beyond", I just mean that we can develop different beliefs and values which depart from the influences of our involuntary affiliations. So, for example, when going beyond what is associated with an involuntary affiliation to *whiteness*, that might mean being affiliated with groups which aim for an awareness of issues centering on racial justice. But because there is no motivation to go beyond whiteness, there just is no difference between their involuntary and voluntary affiliations, making their volitional racism possible. But matters are not always so simple as this, as we cannot always be sure of what one's voluntary affiliations amount to.

²⁴ Though Garcia would likely disagree with this characterization (i.e., that structural forces are the cause the subject's racism), I will mention in a moment why the source does not matter.

The Involuntary Cases

Whereas cases where involuntary and voluntary affiliations make possible overt varieties of racism, what makes possible covert types has less to do with voluntary attachments and more to do with involuntary ones. So, we will need to travel to the underground and consider cases involving involuntary affiliations and unspecified voluntary affiliations. The reason I am labeling voluntary affiliations in these cases as “unspecified” is that it is possible that those who are actively ignorant or implicitly biased are well-intentioned, but it is equally possible that they are not—either they have attachments to anti-racist higher-order beliefs and values, or they don’t. Unless we can read minds, it is best to leave their voluntary affiliations unspecified, because we don’t quite know their status.

There may be some push-back from AI and IB against my “unspecified” label because whites are inevitable generators of racism fed by the fuel of their active ignorance and implicit biases. In other words, perhaps there is *nothing* unspecified about their voluntary affiliations; it may just be the case that they *non-consciously* voluntarily affiliate (as does the volitional racist) with their whiteness, racist baggage in tow. Be that as it may, all I want to say here is that what we can plausibly *assume* is that they are not *necessarily* volitional racists, because they may not believe their prejudicial biases to be true, and they may not be motivated to be unaware.

If we grant this a plausible assumption to make, then how does the affiliation diagnostic system help us understand how AI and IB are made possible via involuntary affiliation?

By zeroing in on an involuntary affiliation to whiteness, we can diagnose how this works. Recall from above that a subject's involuntary affiliations likely influence her psychology. These influences likely betray conscious choice and may produce unnoticed beliefs and values absorbed at a lower order, which in the present case are associated with whiteness (e.g., social capital and/or nonconscious prejudices). Now, if there are no clear voluntary anti-racist or racist affiliations (which we have labeled as "unspecified"), then we can say their behavior is likely determined by their involuntary affiliation to whiteness. So, whatever their voluntary affiliations, the conditions of possibility of AI and IB can be said to be the subject's involuntary affiliations.

1.3.2. Benefits of the Diagnostic System

Let us take stock. In section 1.1 we surveyed some varieties of racism representative of the literature, and in section 1.2 we saw how we might reduce various varieties of racism to some one thing. We have seen the pitfalls of the fragmented (the "many") and reductive (the "one") approaches. The affiliation diagnostic system (ADS) dissolves the dilemma between choosing the many over the one. All I have tried to do is show that the ADS is a *way of thinking* at a second order that makes sense of the whole of the literature by non-reductively unifying different varieties of racism. What I offered is a connective tissue that shows

what makes any type of racism possible. The ADS offers a unifying approach that subsumes reductionists and non-reductionists alike, with no fixed meaning to racism identical in every case.

But allow me to spell out in a little detail the ADS's desirability and benefits. Generally, the benefits of the diagnostic system are that it helps our thinking about racism by offering a meta-theoretical perspective that can unify any type of racism in a commonsense way, a global view of the literature without sacrificing the explanatory power of any one view while being sensitive to their differences. We do not need to undermine, replace, or reduce any view. But let us consider more specifically the desirability of the ADS methodologically and by reference to praxis.

The ADS is methodologically desirable because it offers a framework for thinking that buys us out of the dilemma, a second-order map that enriches and integrates our understanding of racisms.²⁵ We can avoid the undesirable consequences of each horn of the dilemma: possible misidentification of racist phenomena due to narrow overspecialized approaches (horn 1), and the denial of nuance and epistemic privilege (horn 2). The ADS avoids the consequences of horn 1, because it offers a way of thinking that is not confused or mistaken about

²⁵ But what of institutional and structural racisms? One might say that this exists independently of conscious human experience, and thus cannot be diagnosed by the affiliation framework, relying as it does on subject-group connection. We can understand institutional and structural racisms as depending on a network of interrelated involuntary and voluntary affiliations. A first-order theory of structural racism that understands it as a system produced and sustained by white supremacy can be thought through at a second-order by seeing it as conceptually depending on subjects who voluntarily affiliate with groups whose set of beliefs have it that the benefits of structural racism are too good to eschew, and at the same time, subjects who are involuntarily affiliated with whiteness who reap those benefits whether they want to or not.

what we are trying to understand – we know the sorts of people we are talking about and are not mistaking some racists for others. If we know what all the specialized approaches depend on, then we are able to agree on what sorts of affiliations are at stake. And if we know this, we can effectively judge which first-order approach is the correct one for the phenomenon at hand, because we have a simple diagnostic toolkit that illuminates the kind of affiliation(s) that makes the approach and phenomenon possible. It is all conceptualizable in easy-to-understand terms that does not depend on special academic or scientific knowledge, and so it is easily digestible by the public. Anyone can understand what is basic to racism, and anyone can understand why racism takes the various forms it does, those explained in the literature. We see the whole and can communicate it simply and effectively, without the technical and narrow expert testimony of academic sub-fields. This has obvious practical benefits, which I address in the next section, about getting people on board with the findings of the first-order positions.

Another benefit is its capacity to *unify* the varieties of racism, which is desirable because otherwise we may reject or ignore particular varieties of racism as well as the epistemic privilege of non-white philosophers of race. This is due to the ADS's *non-reductive* strategy, which allows each type of racism to explain its target phenomenon. As I showed in section 1.2, there are reasons that a reductive approach centering on one boiled-down explanation is unlikely to yield a be-all and end-all view. Has Glasgow ended the conversation? Can we

pack our things and head home? The point of taking a non-reductive approach is to appreciate that each variety of racism has something important to say and that jettisoning one (or more) in the service of reductive simplicity may leave us high and dry where it comes to explaining some more nuanced, genuine racist phenomena.

The point so far is that the ADS methodologically *retains*. Retention – i.e., preserving the autonomy of any variety of racism – is desirable because we can continue to let each type of racism describe its own phenomena well. The ADS allows us to put them together in a way that does not privilege any one view over another, picking out the right affiliation for the right theory. We should want this, and it matters, because the varieties of racism all explain something different, and we would miss out on aspects of racism were we *not* to retain the autonomy of the disparate views of racism. We get all the explanatory merits from all the positions that we would otherwise miss out on *because* they each explain something different. They can continue to track the moving target of racism, as well as what is new, what has not yet been explored or discovered, and what might change in the future. But beyond unification and retention, the ADS shows that we can think about racisms objectively while avoiding the denial of epistemic privilege, because the ADS centers on the transcendental conditions of possibility of human experience. What could be more central to racism than human experience? Though this may be trivial, the ADS is a solid ground, because anyone could assent to it. So, a white philosopher attempting to

be objective about racism can do so without rejecting or infringing on perspectives she does not have.

Finally, and by way of a brief preview of what follows, the ADS is beneficial vis-à-vis praxis, which here refers to the amelioration of racism, how to lessen it. If we buy into the lesson of the foregoing section, what should we now do if our goal is to lessen racism? Should we pick one amelioration strategy or sample them all? My thinking is that we put the ameliorative strategies together for a complete picture, much like we did for the varieties of racism in this section, where the strategy should fit the problem.

2. Unification to Amelioration

In this section we re-encounter a familiar constellation of themes: the methodological blindness and limitations *à la* the “one” where with the hammer, everything looks like a nail, as well as the narrow overspecialization *à la* the “many” where more is done to complexify the issues instead of simplifying them for the public. The section argues that non-reductively unifying racisms helps orient our thinking about their amelioration, namely that a generalist approach like that of section 1 is applicable to amelioration strategies. The point is that a generalist approach is required *because of* methodological limitations and overspecialization. Resultant public skepticism and blowback is what happens when we cannot talk coherently to the public. So, what the generalist has to offer is a way of thinking practically, a way to gather up the theoretical chaos and speak clearly and coherently before leaping into praxis.

2.1. *Varieties of Racism, Ways of Ameliorating*

There are two moves in this sub-section. The first is to appreciate and acknowledge that the following views are legitimate and have an impact on ameliorating racism. The second is to show that there is always more to the amelioration story that these views cannot tell given how they understand racism. The point is that while no approach to strategizing amelioration is wrong, each is limited by how racism is described. By clearly and coherently defining the problem in such a way as to highlight its wrongs, we can clearly identify how to lessen the problem and *ipso facto* its wrongs. Now, I am not going to offer a first-order all-inclusive amelioration program below, something which “solves it all”. Rather, all I am trying to do is show that if each amelioration strategy is taken as an island, then it cannot fundamentally make racisms better outside that island. Each of these is a tool. Each works within a certain context. What I *am* offering is the idea that if we do not see the whole, we get myopically caught up with each specific amelioration strategy.

Social-structural Racism and Social Amelioration

If we define racism so as to mark off and draw out its socio-political wrongness, then we have a readymade blueprint for combating racism. First, the view locates racial oppression at the social-structural level, revealing the ways it is baked into social institutions which maintain patterns of racial inequality and injustice, shedding light on what might or would be required for mitigating its harms. Second, it expands the concept of racism beyond the individual “bad

apples”, where it is not enough to push the problem off to bad actors within a society.

So, we can take racism in this sense to be a systematized network of oppressive and dehumanizing forces which deprive a racial group of material capital and sometimes life itself while uplifting and maintaining the social position of the dominant oppressor group (Carmichael and Hamilton 1967; Omi and Winant 1994; Harris 1999/2020; Feagin 2006). Some analytical approaches to defining this social-structural concept of racism are prescriptive or normative, with an eye toward criticizing the structure of a society which unjustly disadvantages racial groups (Haslanger 2012, 367, 376, 385-387, Ch. 11; Urquidez 2020, Ch. 7). Whatever our definition and whichever analysis gets us to that definition, what we end up with is a picture according to which racism is, at the outset, a normatively-laden term.²⁶ What this means is that the constraints set on this view of racism lead to assessments on how to ameliorate it which come with their own constraints. That is, if a social-structural explanation of racism is prioritized, we need to come up with solely social-structural solutions to ameliorate it.

We arrive, then, at views of what can be called *social amelioration*, by which I mean attempts at lessening the socio-political and institutional nature of racism. For instance, we might posit that racial progress and the dismantlement of institutionalized racism depends on whites giving up their privileges and

²⁶ Not everyone who is sympathetic to a structural view of racism agrees with a moralistic starting point. See Shelby 2002 and Mills 2003.

eliminating the ways in which they participate in and benefit from “institutionalized patterns of racism” (Stickers 2014, 16). Or perhaps we go for insurrectionism, another strategy which calls for directly challenging a racially unjust status quo to liberate racialized groups (Davis 1983, 1990; Harris 2002; McBride 2013, 2017).²⁷

On these views it makes perfect sense that cutting out the cancer of racist social structures is the *only* way to alleviate racism, because each understands racism *in terms of* social structures and institutions. These views have their virtues, drawing awareness to institutional privileges and challenging the status quo via activist scholarship. Each *goal* is one we should try to reach. But there is a difference between *theoretical goals* and *actual goals*. What I mean by this distinction is that there is a rift between *ideals* to which we aspire and whether the aspirations are *practically feasible*. That is, the *means* by which we reach those goals matter. Things get murky where we consider these goals as actual, as none of them seem practically feasible. Each view sets out farfetched, often idealistic means to achieve the goals of social amelioration, because it seems unlikely that social forces can be overturned, especially if it is true that whites are inevitable beneficiaries of racism. Goals that are morally praiseworthy but fundamentally unattainable do not practically ameliorate racism. Certainly, addressing

²⁷ Or perhaps more radically, we assume Derrick Bell’s “racial realism” thesis (1992), as does Tommie Curry (2007, 145), and consider the possibility that social amelioration is impossible and instead advocate purely for liberation. We might then entertain the possibility that violence against whites is the only way to solve anti-Black racism (134-135). This by-any-means-necessary approach is a strategy about which I am agnostic.

structural racism is an important part of amelioration, but it is not the *only* part of the story. If social amelioration is our only tool, then we limit how other varieties of racism might be lessened.

Racism and AI: Epistemic and Habit Amelioration

How might we start changing minds and not just structures? Look to the content of those minds – what people know and don't know. We should worry about how racism is generated by willfully ignorant whites with bad habits, especially given that AI is structured by social-institutional forces which maintain and manufacture white privilege and supremacy. Coupled with the above, there are reasons that describing racism in this way is valuable. It features how socio-cultural forces blind whites to their role in furthering structural racism. But importantly, it also shows that whites *welcome* this blindness (as a matter of conscious choice or not), because it serves the purpose of upholding a racist status quo that whites are motivated *not* to recognize or acknowledge in any meaningful way (Mills 1997, 2007). So, our first step should be getting *aware* of structured ignorance and mending bad habits (Sullivan 2006; Hoagland 2007; Outlaw 2007; MacMullan 2009; Medina 2013; Al-Saji 2014).

So, we get models of *epistemic* and *habit amelioration*, by which I mean attempts at lessening racism as generated by AI by improving knowledge practices and modifying habits for the better. For instance, we might argue for developing epistemic virtues such as humility, curiosity, and open-mindedness (Medina 2013, 42). Or perhaps we should appreciate and recognize *relationality* to

see the *interdependence* between oppressor and oppressed to open better lines of communication (Hoagland 2007). Still more, education that challenges the dominant (white) perspective and mitigates ignorance may also be a necessary strategy (Outlaw 2007). Any which way *epistemic amelioration* is spelled out, we may begin to override certain ways of actively ignoring the status of the oppressed.

Yet, beyond *epistemic amelioration*, white habits may still get in the way of progress where AI is concerned. So, we may offer views of *habit amelioration*, strategies which emphasize the modification of, or resistance to, habits of white privilege (Sullivan 2006) and habits of whiteness (MacMullan 2009), or the interruption of racializing habits of “seeing” (Al-Saji 2014). We may, then, begin to counteract the maintenance of a white homogenized worldview. On these views, addressing bad epistemic practices and habits is a way to lesson racism, because they understand racism in terms of (a willful lack of) knowledge and bad racialized habits. These views seem more realistic than the means for *social amelioration*. They are helpful for shining a light on the ways whites are unwittingly racist and provide tools for resisting racism that results from AI. They are more practically realizable given their focus on explicit strategies people can adopt.

Yet, alone, this too is not enough because there are people who know better, care, and manage their racialized habits yet still demonstrate bias as the Implicit Association Test shows us (Kelly and Roedder 2008). That is, IB, which is

not a matter of knowledge or racialized habits, is left on the table, another set of people otherwise un-ameliorated if we view matters only through the lens of AI.

Racism and IB: Trained Amelioration

Let us, then, turn to IB. If we understand racism (personal or structural) as a result of people's implicit biases, then again, we have a handy readymade blueprint for lessening it: thwart the causes of implicit biases in the brain. There are good reasons to describe racism as the result of IB. It does not make racism out to be a wholly structural phenomenon, while at the same time appreciating the force that socio-cultural racism has on individuals' psychology. It opens our eyes to the complicated matter of nonconscious racism.

So, it makes sense that amelioration here would involve training out people's biases, which I am calling *trained amelioration*. There are several approaches for doing so. For instance, one strategy is to "break the habit" of implicit bias by intervening on subjects so as to train them to become aware of their automatic bias (Devine et al. 2012). Another is to cultivate "positive personal contact" with members of an out-group which a study shows reduces racial bias in police officers (Peruche and Plant 2006). Two further, competing routes are on the one hand, to disaffirm stereotypes (Kawakami et al. 2000), and on the other, to affirm counterstereotypes (Gawronski et al. 2008). The first route finds that saying "no" to stereotypes with repeated practice reduced automatic stereotyping. The second finds that saying "yes" to counterstereotypes after training reduces the activation of stereotyping. On these views, to alleviate bias is

to alleviate racism, because racism is understood as a kind of bias. So, we have strategies that are more realistic for those who already know better, and helpful for correcting people's instant biased responses.

2.2. A Generalist's Apology

Each of the amelioration strategies above lessen their target racisms *within* their purview. Each is a tool that works within a certain context. Use the hammer with the nail, the saw with the plank. But the point I want to make at this juncture is that none lessen racism *outside* their purview, that if we do not see the whole, it is likely that we will get caught up with each specific strategy and each preferred description. With my hammer, everything looks like a nail; with my saw, it all looks like planks. Still more, these parts alone do not buy everything where it comes to ameliorating racism as a whole, and on their own, they may be unconvincing to skeptics, the very people we *want* to convince, our targets in the public. So, we need a generalist approach. We need to put ameliorations together coherently so that we get a clear, general picture of how to think through mitigating the whole of racism. But there is an immediate, obvious objection to what I am saying. Given that piecemeal strategies ameliorate within their purview, there is no need for a wholistic picture, no need for a master planner. Each can operate on its own, within its purview, and chew away at various aspects of the racist whole. So, we don't need a unifying order. We don't need a whole.

This much I concede. But the point is that the narrow specialization of these strategies alienates the public and deepens their skepticism about racism. A case in point is the recent demonization of Critical Race Theory.²⁸ Again, as in section 1.1., we need to ask: who are these descriptions of racism and amelioration strategies *for*? Are they for *me*, an academic; are they for *you*, likely another academic? All I am driving at is that the narrower our descriptions and strategies get, the more extreme they get in their complexity. And the more divorced from everyday experience they get, the easier it is for the public to be skeptical. We are not trying to convince people who are already sympathetic; if we want to sell a progressive liberal on the nuances of racism, then we might give them a copy of Medina (2013). But I would venture that no one we are *trying* to convince will say they “signed” the Racial Contract *à la* Mills (1997). We won’t sell a racist cop on AI, and IB would likely be an uphill battle. A skeptical, resistant public will likely have no patience for any of this, they will likely think it is academic hogwash cooked up by eggheads, and not bother to understand it. The idea is that what works for some might not work for others. But if we take the time to talk about it, generally, they may come to understand it.

So, we need a generalist approach. As I said earlier, the philosopher’s role should be a functionary one *for the public*. We should elucidate, not complicate. The best moral goals in the world don’t matter if they are unconvincing and likely to never happen. To reiterate what I asked in 1.1., what could be more

²⁸ For some overviews and examples of what I am referring to, see Butcher and Gonzalez (2020), Krasne (2020), Goldberg (2021), Harris (2021), and Sawchuck (2021).

ameliorative than something the public could understand and get on board with? A cross-specialization generalist can coherently create a bigger picture, because with a big picture, we can clearly identify our phenomena, our audience, and our amelioration strategy. This is praxis at its best. It matters *how you present it*. All the meta-theoretical straightening out now matters.

So, the point I am driving home is this: *let's get things clear descriptively in order to make things better*, because the way we *describe* can either limit or unshackle the way we *prescribe*. We shouldn't obfuscate moral issues and practical ones. The former are aspirational goals. The latter are a *process*. Now, I am not suggesting that we divorce descriptive work from a moral context. How could we, given that our target is racism? Rather, what I *am* suggesting is that we should not let moralistic upshots and programs *define* our descriptive work. Letting any one of the above approaches determine how we capture racism not only misses different kinds of racisms, but puts blinders around our eyes, keeping us from seeing how to make racism as such better. So, not doing neutral descriptive work leads to consequences that are not as good as they could be for racism itself, just for particular varieties of racism. The reason thinking through the theoretical, descriptive apparatus is important is that it gives us practical purchase; it enables us to identify how to lessen racism as a whole. And none of this must be mysterious. I showed that we are not, and need not be, limited by our descriptions by putting racisms together where the formal ground is the way humans affiliate with each other. Putting together a range of racisms by

affiliation allows us to see how we can put together a range of ameliorative strategies. No longer are we limited or constrained by how we understand particular varieties of racism; we can see it as a whole. We should care about a “sample” approach because the consequences are what matter most – the right ameliorative tool for the right racist phenomenon. The consequences of weaving together racisms are the ability to weave together ways of ameliorating it, a combined picture of racism with a combined picture of amelioration helps us see how to pick apart the practical problems we confront.

Conclusion

If different varieties of racism are ameliorable, then they are avoidable, and if they are avoidable, then we can begin thinking about who or what is responsible for racism. But this is tricky, because one’s thinking about responsibility may well depend on one’s preferred description of racism and its purported wrongs. Should we think about racism and responsibility in a personal or socio-political sense? Should we go narrow or go broad? As luck would have it, we have a set stage for thinking this through, because there are individualized senses of racism and of responsibility, and social senses of racism and of responsibility. So, if we choose to go narrow and focus on individual racism, then we can make use of an individualistic account of responsibility to say that so-and-so is *personally* responsible for their racism. But we could choose to go broad, focusing on social racism, and make use of a social account of responsibility to make a case for *social* responsibility for racism. So, should we

pick one over the other? Should we emphasize one over the other? What is the cost of doing so?

My thinking is that we will run into a familiar hitch: if you prefer a social understanding of racism, you will likely prefer a social/civic sense of responsibility. If you prefer a personal understanding of racism, you will likely prefer an individualistic sense of responsibility.²⁹ If we hunker down in either camp, then we create responsibility loopholes. If we locate responsibility at the individual level, then so long as one is not actively racist, then one no longer has any accountability outside oneself. But if we locate responsibility at the social level, then one can evade responsibility by claiming it is unfair to be held to account for an unchosen participation in systems of oppression that one did not “sign up for”.

My thinking is that *before* asking who or what is responsible for racism, we should *first* re-frame the very idea of responsibility – the topic of chapter 4. What we have seen so far is that there is no straightforward way of locating responsibility in the sphere of individuals or society without losing out on one or the other. So, there is something to learn from racism which makes it a kind of methodology, an approach for thinking through other philosophical problems like that of responsibility. The next chapter will provide some headway towards thinking through how the personal and the social-structural intersect – that

²⁹ Where and why do such preferences arise? Although I will not discuss this further, I would wager that they have something to do with one’s attitude towards what can be done in terms of amelioration and hence what we are responsible for.

social-structures *play out* in the individual's mind – the nexus point of the two.

This will, in turn, provide headway into answering how we ought to re-think responsibility, doing away with the binary between the social and the individual.

Chapter 3 *Overlooking: It Could Happen to You*

Introduction

“Yes, all whites are racist, but not *me*.” Is this what any philosophically educated white person in the area of race and racism would say? Probably not. Is this what they might assume about themselves (consciously or not)? Perhaps so. But why might they think so, and how might this psychological tension be described? This chapter targets whites of this ilk by answering that question, identifying a novel phenomenon of racism that I will call *overlooking*. These are whites who know better and have anti-racist values, who do not seem to be ignorant (actively or otherwise), nor are they acting in racist ways as a pervasive pattern. Whites who know better and are not usually doing racist things may, consciously or not, feel as if they are not part of whiteness as addressed by active ignorance or implicit bias. In a similar vein to the phrase “I’m not a racist, but...” such persons might suppose “I’m white, but...” This chapter does a bit of performative amelioration, then, for the overlooker who might lurk in me, the author, or (if the reader is white¹) might lurk in you, too. Once we know about it, we can take a hard look at ourselves, and perhaps prevent it.

Though this chapter’s first goal is to contribute a description of a novel *phenomenon* of racism to the literature, the phenomenon is continuous with the

¹ Of course, not all racists are white. So, what I have to say would hold for any racist person. But my reason for singling out whites is that in terms of asymmetrical power, the effect of white racism in terms of its harm is greater than other racisms perpetuated by other racial groups. See Blum (2002a, 33-52) for a discussion of this point.

legacy of the social-structural phenomena of active ignorance (AI) and the psychological phenomena of implicit bias (IB). To demonstrate novelty, I will argue why overlooking is not reducible to AI and IB, while noting continuity with the sorts of racist phenomena those positions model. I will show that overlooking continues what the AI and IB approaches have already started, a further analysis of unintentional racist phenomena. It is not that I favor ignoring social structures over individual psychology, but that these social forces *play out in our minds*, where the social *is* the mind and the mind the social, a picture which is mapped out by the previous chapter's *affiliation diagnostic system*. Indeed, it would be hard to imagine a contemporarily relevant account that denied social structures, which might imply that we are all "done" with racism, hence tending to contribute to and perpetuate it. Just as IB's psychological lens shows how social forces lead to whites *as such* being non-consciously biased, overlooking offers a more psychologically complex manifestation of these same social forces – showing us how individuals might non-consciously think they are *personally* done with racism, given the right amount of knowledge and care. Both senses generate smooth sailing for racial prejudice. Thus, while overlooking is methodologically continuous with the ethos of AI and IB, it offers a new vein of complexity with clear ramifications for those who think deeply about racism.

The primary aim of the chapter is to motivate a new conversation about *personal racism*, that our conversation about racism should not depart entirely from a focus on the individual. While I will mention some descriptive

consequences of the chapter, I emphasize the moral and practical consequences of overlooking. Specifically, I draw attention to *moral evasion* – escaping moral accountability by appealing to exculpatory conditions – and argue that unlike AI and IB, there is nothing in overlooking which implies evasion, either by shifting blame to social-structural forces or automatic brain processing. In the conclusion I highlight some undesirable practical implications of AI and IB. As it is psychologically unrealistic to be in a constant state of shame and vigilance, I argue that such views likely lapse into the sort of risky thinking that one has arrived on the side of righteousness, that one is *personally* done with racism even if whites in general are not. This implies an artificial division between those who are “in the know” and those who are not, which has the potential to proliferate unintentional racisms. I contend that overlooking dissolves this artificial division and shows that no white person can claim to comfortably roost on the “in the know” side of this division.

1. The Overlooking Phenomenon

Consider Kim, who is a white professor of critical race theory. She is an expert in the field, having devoted her entire career to its study. Her particular research emphasis is racism in all its various guises. Clearly, Kim is knowledgeable about such matters. When confronted time and again with difficult and uncomfortable truths about her participation in, and benefitting from, social networks of white privilege and supremacy at the expense of people of color, she is not so fragile as to back down from recognizing her role in that

social situation. She does not resist knowing more about her complicity. Though she has never herself experienced racism, she has nonetheless devoted her life to understanding it, speaking and writing about it, doing consultant work with area businesses, and to engaging in public demonstrations seeking to raise awareness about it. She's been arrested at demonstrations, most recently during the summer of 2020 while marching at a Black Lives Matter protest. She takes seriously what she's come to know; she cares.

But while she devotes herself to study, and to antiracist activism, she does not attempt to delude herself into thinking that she has "arrived" at some antiracist or non-racist ideal. Aside from people of color, who would know and care more than Kim?² But she hasn't gone through life without some near misses and embarrassing slipups. Let's say that Kim has just watched a documentary about the antebellum South. Shortly thereafter, Kim helps her Black colleague carry heavy boxes of books to a new office and begins whistling the opening note of "Swing Low, Sweet Chariot" (a song featured in the documentary). She stops herself, realizing all in a flash that whistling the tune would be to thoroughly ignore the history of slavery associated with the spiritual that could be offensive and cause harm to her black colleague. Now, this is not something that always happens to Kim. She's not always slipping up, and she does stop herself from

² The first clause of this sentence is meant to show my commitment to subjective perspectives laid out in chapter 1. Kim has no *complete* first-hand experience or experience-based knowledge of what it is like to be a person of color, though of course this does not rule out that she can analogize her own possible experiences of other forms oppression to that of racism.

whistling the full melody. Nevertheless, Kim is not perfect, and she sometimes overlooks what she would otherwise not choose to do.³

A case like Kim's can happen to anyone. You might recall doing something like the above or know someone who did. You probably know someone like Kim. *You* might be Kim. We all know of embarrassing cases like this. What the example suggests is that the right kinds of knowledge, values, affiliations, and/or experiences are not prophylactics against the phenomenon. Kim knows better and cares, but it makes no difference. She nevertheless overlooks. The example shows, *prima facie*, that overlooking is accidental, that despite knowing better and caring about something, one can do what one otherwise would not consciously choose to do.

Other examples besides the case of Kim can be offered. For instance, consider a corporate drone who has personal experience with races other than his own – e.g., he has learned about issues from friends, grew up and lives in a racially diverse community, and is affiliated with people of color. Let's say he has an attached "wokeness". He has a personal commitment to antiracism and strong progressive leanings. He cares. In any event, he works in a skyscraper, and he's annoyed at having to take a long elevator ride to the ground floor so he can go outside to smoke. On one elevator ride down, he says, "If only it were the 50's again," gesturing to his cigarettes, despite knowing what the decade was

³ Other examples of overlooking could have different degrees of harm. They could also have different degrees of knowledge, care, control, and awareness. The case of Kim is but one example of the phenomenon.

like for black people.⁴ Some of his black colleagues are in the elevator. He becomes aware of his action all in a flash.

Or consider a software program developer who has an experienced understanding of structural racism and is by all measures well-informed. Like the smoker, she's a committed antiracist. Suppose she is tasked with designing a healthcare algorithm. She designs the algorithm for health institutions to determine access to special healthcare based on need for such access (e.g., high-risk illnesses). She uses "costs" as a stand-in for "need based on illness". Once set into motion, the algorithm contains a critical bias against black patients, since significantly less is spent on this group in comparison to whites.⁵ As a result, black patients are not selected as needing special care – they *appear* to be of better health than just-as-sick whites. She becomes appallingly aware of her action's import sometime later.

One should wonder how much the overlooking phenomenon occurs. While I offer other examples besides Kim, I focus specifically on her case because scenarios like it are, I'd wager, ubiquitous. The phenomenon hits many whites where they live – those who think they have put in the work, that they are knowledgeable and care. Yet, they nevertheless watch their racism unfold before their very eyes. There is nothing eccentric about overlooking, and if it is ubiquitous, then we should focus our attention on it. *I* have overlooked, and

⁴ Let's say he's been binge-watching the show *Mad Men* lately.

⁵ See Obermeyer et al. (2019) and Evans et al. (2020) from whom I draw inspiration for this example.

perhaps *you* have, too. Overlooking is meant to demonstrate an “insider’s” account of white racism, not by appeal to notions such as white guilt or fragility, but in a way that is well-grounded in the white perspective – what it feels like to participate in whiteness from the point of view of the white person.

1.1. *What is Overlooking?*

The overlooking phenomenon is a species of what I am calling *unintentional racism*.⁶ By “unintentional racism” I mean racism which is not done on purpose or deliberately, with no aim in mind, and involving no willing or desiring. It is accidental. The example illustrates that Kim has no motivation towards racist actions. Given her epistemic competencies and values, she would not have acted as she did if she had consciously chosen her action.

Just so we are on the same page, I am going to formalize this as a matter of clarity. Just as typos are kinds of mistakes where one fails to enact what one knows and values (e.g., proper grammar), the same goes for overlooking. By “overlooking” I mean a psychological mishap, where when a subject *S* overlooks: *S* (a) unintentionally fails to remember facts about the world and remember something she knows about the world which results in (b) *S* performing action φ due to forgetfulness over and against possible action ψ despite *S*’s otherwise knowing that ψ was the more appropriate action given her value scheme. By “psychological mishap” I mean a cognitive lapse where an otherwise informed

⁶ There may be other phenomena of unintentional racism beyond overlooking, but I will not explore them in this chapter. Some racist events caused by active ignorance and implicit bias arguably fit under the label *unintentional racism* as well, though I will not defend that claim here.

subject acts in a way she would not have if she were presently mindful of the circumstantially relevant information.⁷ By “forgetfulness” I simply mean failing to remember something. Here, when subjects forget, they fail to remember something they otherwise know. Subjects are liable to make unintentional mistakes even when they are informed. By “informed” I mean being aware of (i.e., noticing) and knowing some state of affairs. I draw no sharp distinction between “maximally” and “minimally” informed. While this may be a way to explain the differences between experts and non-experts, the point is that for our purposes here, subjects are at least minimally knowledgeable.⁸

2. Active Ignorance, Implicit Bias, and Overlooking: Continuities and Departures

Two existing models seemingly prepared to handle the example of Kim are active ignorance (AI) and implicit bias (IB). Whereas one might expect that IB is more promising than AI, I will show that it too will not work. The phenomenon is not modeled well by either, because overlookers are not ignorant (active or otherwise⁹) and have some conscious control over their actions. By “model” I mean that these approaches do not apply well to the example, rather than claiming that they are just wrong *in toto*. Yet, overlooking is not reducible to

⁷ This is similar to *akrasia*. But, overlooking is not *akrasia*. *Akrasia* involves conscious awareness—more strongly, *knowledge*—of what one’s doing and lacking the will to do otherwise, whereas overlookers may *become* aware, and sometimes *have* the will to do otherwise.

⁸ This point about minimally knowledgeable subjects is akin to the “thesis of cognitive minimums” (Medina 2013, 127ff).

⁹ I write “active or otherwise” to express the distinction between *active* and *accidental* ignorance as drawn by Bailey (2007, 77). The former applies to deliberate refusals to know, while the latter applies to the ordinary sense of ignorance where this means a lack of knowledge which can be filled in.

AI or IB, because it does not match the explanatory approaches or assumptions of either model. These two positions only tell half the story of overlooking because they do not explain the *conflict* which Kim so obviously experiences. Thus, while there are *some* aspects of the overlooking phenomenon for which AI and IB can provide clarity, I demonstrate that the example is truly a novel phenomenon of racism.

2.1. Active Ignorance and Overlooking

Suppose that upon immediate assessment we are inclined to say overlooking seems to be a form of ignorance, that it is either explainable by or reducible to it. One conventional approach taken by theorists of AI involves a point about white resistance to knowledge of racism in any of its forms (Mills 2007; Bailey 2007; Medina 2013). Supplementarily, some approaches focus on a lack of motivation *to* know that whiteness is a maintenance of a racist status quo (Mills 1997; Sullivan 2006; Alcoff 2007; DiAngelo 2011, 2018). Whether as a means to maintain one's social position and capital or as a function of bad epistemic practices, it is perhaps the case that overlooking is just a token of the AI type.

In this spirit, we might say that, generally, overlooking is explainable by the structural forces that maintain white ignorance (e.g., *white privilege* and *white supremacy*) and is thereby another instance of the "epistemology of ignorance" (Mills 1997, 2007). The basic approach here would be to claim that white folks, globally, are all ignorant to some extent, in which case AI could explain overlooking. Kim is just a token of an ignorant white type, and so it is in virtue of

her whiteness that she succumbed to the forces of white ignorance. It is therefore not enough to say, by caveat, that Kim is not ignorant.

Let us build the case further. Charles Mills (1997), for instance, explains that whites are socially motivated to refuse their own state of ignorance. They are *necessarily* ignorant that they are ignorant. Through the creation of the *Racial Contract*, whites have effectively blinded themselves. White ignorance is deliberately designed and enacted to enable participation and benefit from racialized social inequalities. Thus, even where there is clear evidence that racial injustices exist, actively ignorant subjects are motivated *not to know* so as to maintain their privileged social position. So, that overlookers are not “presently mindful of the circumstantially relevant information” could be taken to mean that they have non-consciously succumbed to the normative force of the Racial Contract, a “structured blindness” to anything related to race (19). Despite their knowledge and care, overlookers cannot escape the “cognitive dysfunctions” (18), or “difficulties” (93), generated by the principles prescribed in the Contract. The dysfunctional thinking identified by Mills is paradoxically *functional* insofar as it serves a psychological and social function of reinforcing the validity of a white homogenized worldview. The overlookers’ inability to think from perspectives not their own is also dysfunctional in a broader rational sense – of validating an irrational racist worldview – but, additionally, serves a normative function of valuing that worldview to the extent where others are sometimes not known to exist.

In cases of overlooking, however, there is no blindness to others' perspectives *as a pervasive pattern*. Rather than "global cognitive dysfunctions", overlooking involves cases of "local" or "episodic" cognitive dysfunction. For the overlooker, the dysfunction that Mills describes serves no paradoxical functional purpose whatsoever because she does not seek homogenization or marginalization of the perspective of other races. That value is something to which the overlooker would not subscribe, and yet her actions betray exactly that value. This indicates that at some level this social normative force is operative in the overlooker, even though it is contrary to a broader epistemic and value system she employs in most cases. Thus, it may be true that Kim cannot escape these forces, but that does not entail that she cannot *ever* resist them. Is the person who has committed her life to studying race and racism *utterly* ignorant, with no control over her actions? It seems more plausible that folks like Kim *do* know better but are *still* susceptible to racism as generated by the kind of global white ignorance theorized by epistemologists of active ignorance.¹⁰

Be that as it may, we could entertain more specific routes that might reduce overlooking to AI, namely that: (1) overlooking is a function of bad epistemic practices and a lack of motivation to know better, or (2) there is some ignorance in the moment of overlooking.¹¹ I will take up each of these routes and show that neither entirely pass muster.

¹⁰ This, of course, is not to say that there are not totally ignorant whites in the sense that Mills identifies.

¹¹ We might construe forgetting as a kind of ignorance, that overlookers are motivated to forget what they already know. Perhaps they are "motivated" by the "habits of white privilege"

Perhaps overlooking can be captured as a mode of resistance for the protection of white privilege and white social capital (Medina 2013) – the result of some epistemic vice revealing the overlooker’s avoidance or un-acknowledgement of any evidence that may threaten their socially privileged position. They may be, in José Medina’s terminology, *epistemically arrogant*, thinking that (whether they are minimally or maximally informed) they have arrived at “enough” knowledge about race, racism, and structural oppression that they need not try to know more about the perspectives of those who are oppressed (30-33). This would explain the accidental nature of overlooking, that taking oneself to already know better and to care enough eliminates any need to be vigilant of possible racism. And so, it would seem that overlooking can be handled by appeal to epistemic arrogance, and thus reducible to AI.

But how could the foregoing be true, given that overlookers *want* to know and *care* about knowing? As a critical race theorist, Kim has devoted her life to understanding race and racism. It would be unlikely that she does not care at all. If it is true that whites like Kim can never genuinely know how racism operates or her position in a privilege-domination-oppression nexus, then why would she continue to seek understanding? It could only, then, be that she really cares about knowing better. Kim has clear attachments to antiracist values, and whether it is by education, experience, or affiliation, she is informed and

operating as un- and sub-conscious defense mechanisms serving the maintenance of white privilege (Sullivan 2006) or by their “white fragility” (DiAngelo 2011, 2018), and so they have a tendency to forget that which is possibly racist from time to time. But, unless Plato’s recollection theory of knowledge is correct, this would be absurd; one cannot forget what one does not know.

motivated to be informed.¹² There is neither a refusal to know or acknowledge racial injustices and structural oppression (Bailey 2007; DiAngelo 2011, 2018), nor an encompassing closed-mindedness (Medina 2013, 35-36), nor a blindness to a world which uplifts her position at the expense of people of color (Sullivan 2006). Thus, unless forgetting is a bad epistemic practice, which is unlikely, this route seems to encounter a dead end.

Yet perhaps we should not be too quick to proclaim that Kim is not arrogant. She may not be shot through with arrogance, nor even usually arrogant. But maybe *some* part of Kim *is* arrogant, she thinks she *is* done and does not need to be on the lookout. While she may not be *pervasively* epistemically arrogant, it could be the case that she is sometimes arrogant in the moment of overlooking. This takes us to the claim that ignorance is perhaps operative *in the moment of overlooking*. A momentary not-knowing is something anyone experiences. So, let us suppose that in addition to epistemic arrogance, “ignorance in the moment” means something like George Yancy’s (2008) notion that whiteness can sometimes “ambush” even well-meaning, antiracist white people, where this means “being snared and trapped unexpectedly” by one’s whiteness (229). On this supposition, it might just be that an overlooker takes herself to have “arrived” at the antiracist ideal and suddenly (in the moment)

¹²Of course, knowledge is not *sufficient* without values – knowledgeable sociopaths exist, after all. For overlookers, it is only because they are *motivated* by what they value to know better and act accordingly. Yet, what overlooking presents is how that link between knowing better and acting accordingly breaks down.

“undergoes a surprise attack” (229), perhaps by nearly whistling “Swing Low, Sweet Chariot” while helping a black colleague lift some heavy boxes.¹³

Now, the foregoing phenomena are all valid; I do not reject that they occur. But even if these occurrences describe the overlooker in the moment, they do not at all explain the *conflict* that happens in the overlooker’s mind – i.e., the reasons *why* she sometimes does not act in accordance with her beliefs and values – and so they only tell half the story of the phenomenon. All these phenomena can do is describe certain environmental conditions which make “slipups” possible. They are kinds of external determinants which say nothing about what is happening in the mind itself, why Kim has the conflict she has. Why is Kim on one day being ambushed, but on the next day isn’t? There is nothing specific to overlooking that suggests these things are always happening as a pervasive pattern, a point which will continue to bear fruit below in the critique of IB. On the contrary, that Kim is who she is suggests that it is probable that she is vigilant about ambushes. What is more is that there is nothing in the example or description of overlooking suggesting that subjects have, or *consciously* take themselves to have, arrived at an antiracist end. If the “surprise attack” occurs only when subjects consciously consider themselves to have

¹³ Perhaps overlooking “in the moment” might also be partially captured by Sullivan’s (2006) “ontological expansiveness” (10), where not having been harmed by stereotypes, prejudices, or racism generates an ease of connection to one’s environment with carefree feelings and no concerns aside from oneself into which white folk slide. These experiences could create the conditions where, in the moment of overlooking, overlookers are in a state of momentary ignorance, effectively not knowing what they are doing.

arrived at an antiracist end, then it is unlikely that overlookers are falling prey to an ambush.

Although it may not be a pervasive part of her behavior (i.e., she is not *always* being ambushed), maybe Kim non-consciously believes that she has arrived at an antiracist end, and that she already knows enough. Some part of her may simply think she is not susceptible to the features of active ignorance Medina and Yancy address. This seems more sophisticated than asserting that Kim is simply a token of the normative force of whiteness. There is no getting around the social-structural processes of white privilege and supremacy, and these processes bleed into the minds of even the most “virtuous” of white folk. So, while overlooking is not reducible to AI, the position helps to show us just how pernicious social-structural forces can be, that even a subject such as Kim can succumb to that which she explicitly disavows. Yet, none of what is characteristic of AI explains the inner conflict in Kim’s mind – doing what she otherwise would not have done. Thus, overlooking is not just a form of AI, contrary to our immediate assessment. So, we are left with Kim’s internal psychological conflict unexplained, something which *ought to be* explained. Perhaps IB can complete the story by reducing Kim’s mishap to her automatic brain processing. Though she is not wantonly racist and sometimes exercises control over her actions, we might say IB deftly handles those instances where things go wrong for Kim, a question to which I will now turn.

2.2. Implicit Bias and Overlooking

Even if AI does not tell the entire story of overlooking, IB may yet fill in the details. But before we get to why overlooking may be reducible to IB, let us begin with some preliminaries. The conventional approach to IB is underpinned by dual-system theory, an empirical psychological account of information-processing where the mind consists of two systems responsible for two processes (Stanovich and West 2000; Kahneman 2011). System 1 is a fast, automatic, and involuntary processing which is prone to bias; System 2 is a slow, deliberative, and voluntary processing that is rule-governed (Kahneman, 20-24). IB relies solely on System 1, which will have serious consequences for its capacity to capture overlooking.

Perhaps it could be argued that (1) the failure to enact one's knowledge and values yielding unintentionally racist acts is just a different way of saying that anyone at all can succumb to biases outside their control. We might also say that (2) the overlooker's "psychological mishap" is explainable by automatic System 1 processing in the mind of which she is unaware. Though it would seem that IB neatly overlaps with overlooking such that the phenomenon is not novel, I will argue that neither of these claims are all right, thus showing that overlooking is irreducible to IB.

It may be the case that the failure to enact knowledge and values is just a different way of saying that overlookers did or could not exercise any control over their biases. So, their mishaps seem easily explainable: their failure to enact

what they know and care about is simply a function of the uncontrolled operation of bias, an automatic System 1 process. But is it plausible that Kim could *never* stop herself from acting in biased ways? If not, then what is it, *within Kim*, that allows her some measure of control over her behavior? And what could “some measure of control” mean? There must be something preventing her from *always* being biased. But for IB, there just is not an explicit empirical mechanism that would explain this capacity in Kim. There is some conceptual *acknowledgement* of control: that there is *usually* none (Kahneman 2011), some *pro-active* control (Amodio and Swencionis 2018), or that biases cannot be controlled *directly* (Holroyd 2012), where “implicit” usually implies a lack of conscious control (Hardin and Banaji 2013). But what *explains* Kim’s control? Curiously, there is nothing at all in IB to do so, but the empirical foundations of IB, dual-system theory, offers an easily understandable explanation: *System 2 regulates System 1*.¹⁴ In the same way, for example, I might be prone to overestimating the number of cars on the road that are like my own (an example of the “availability bias”; Kahneman 2011, 131-135, 425-427), I can stop (i.e., I can regulate) my automatic thinking by being aware of my own humanly biased tendencies – “I just think that, because I’m on the lookout for that car, and not because there

¹⁴ Now, the automatic-System 1/controlled-System 2 binary (see Frankish 2010, 922) that is foundational to the dual-system theory upon which IB relies (Banaji and Greenwald 2016, 54-58) suggests that we can, over time, significantly reduce bias by “breaking the habit” of bias (Devine et al. 2012), perhaps through bias training (Correll et al. 2007; Gawronski et al. 2008), or “positive personal contact” with out-group members (Peruche and Plant 2006). But whichever way we slice control, some kind of external pressure is applied (i.e., training), and this externality, as I will explain shortly, is the crucial differentiator to overlooking.

really are more of that kind of car.” The crucial difference between this level of regulative control and IB is that *there is no internal higher-order ability* to prevent oneself from doing or saying stereotypical things. One just, for instance, takes an anti-IB training, and then (presumably?) one cuts the bias off at the source. Is one then done with racism? This seems implausible, and worse yet, *implies* that those who are trained might take it to mean that they are all through with racism, stereotyping, prejudice, and so forth. This is the sort of thinking that might let us “off the moral hook”, as it were. Yet the import of overlooking is that we are *never* finished.

How is Kim’s capacity to prevent bias different than IB’s anti-bias training? The question I am pushing here is not whether there is control or not, but instead, where is control *located*? Kim’s “regulation” differs from IB training because the latter’s “regulation” is *external*. One relies on others to help one control one’s actions and automatic thinking, instead of taking a serious look at oneself. The difference is precisely that Kim is capable of *internal* regulation. Training System 1 (for IB) is about conditioning a certain stimulus with a certain response, involving no System 2 processes whatsoever (Kawakami et al. 2000; Blair et al. 2001; Dasgupta & Greenwald 2001; Devine et al. 2012). The general idea is that post-conditioning, subjects automatically react “correctly” every (or most of the) time. But for the overlooker, System 2 actively, deliberately controls (i.e., regulates) System 1 responses, rather than exercising conditioned

mechanical, rote responses.¹⁵ In Kantian language, control for the externally conditioned on IB is *heteronomous*, whereas for the overlooker, control is *autonomous*.¹⁶

The chief descriptive insight of dual-system processing is not only that agents succumb to bias, but that they may also process information rationally, according to their higher-order beliefs and values. The chief normative insight, as seen in second-generation (Thaler and Sunstein 2008) dual-system researchers, is that agents *ought to* regulate their tendency towards biases by engaging System 2. That is, *anyone* can regulate their automatic impulses, provided they have the right higher-order beliefs and values to do so, though no one can regulate perfectly all the time. It is therefore implausible to suppose that Kim has no internal control over her own actions, and if she does have internal control, then her overlooking is not reducible to IB. Kim's System 2 regulation is clearly part her story; we cannot simply say that automatic System 1 processing is always what is responsible for Kim's behaviors and actions. If that were the case, Kim would not just make the occasional error, but instead would make *systematic predictable errors* (Kahneman 2011, 270), acting on every racist impulse that "pops" into her head. But once one has a properly functioning System 2 (i.e., one knows better and is able to regulate one's automatic responses), one then does

¹⁵ This is not even to mention that training out biases does not necessarily *entail* that subjects arrive at new or revised beliefs and values such as, for example, a belief in the value of anti-racism and a motivation to care about it.

¹⁶ For example, IB seminars for police officers involve an external party training the officers who need to be "enlightened" (Peruche & Plant 2006; Correll et al. 2007). But the spirit of the overlooker is that she does not need others to do the work for her.

not always act on those biases even if they still live in the mind as heuristics. As was the case with internal *conflict* (unexplained by AI) above, this internal System 2 *regulation* (unexplained by IB) is an element of overlooking that ought to be explained for a developed explanation of the phenomenon.

The overlooker's internal regulation thus belies the general construal of IB, where race is concerned, as an automatic, involuntary prejudicially biased response. Though even well-intentioned, self-described antiracists are susceptible to such responses (Kelly and Roedder 2008, 525-526), as in the "ambush" phenomenon presented above, given the capacity for internal regulation, overlookers are not prejudiced beyond all control, and do not generally act prejudicially. It is not farfetched to presume that someone like Kim, a white professor of race and racism, is probably not *always* biased against people of color. Were Kim simply implicitly biased, then she would behave as such either all of the time, or at a minimum, most of the time. While there are certainly people who are likely to do this, this does not describe Kim. Rather, she is susceptible (as we all are) to making these sorts of errors. Similar to the foregoing critique of AI, there is a *lack of a pervasive pattern* of biased behavior. So, assuming a view where everyone is determined by their biases to act prejudicially is eluded by overlookers: they neither consciously undertake racist actions, nor do they have no conscious control over them. The partial control here just means that sometimes System 2 works and sometimes it does not – for overlookers the agential "light switch" is neither *just* on nor *just* off, by which I mean there is

neither complete control nor no control at all. More often than not, overlookers have the capacity to do otherwise. They are free – minimally in the sense that they could have done otherwise – rather than always (implicitly) mechanically determined by bias. It would seem a plausible assumption to make that someone like Kim retains this partial control emblematic of overlooking.

Yet, even if overlookers exhibit no pervasive pattern of uncontrolled biased behavior, it could be argued that overlookers' psychological mishaps are reducible to System 1 processing, where "mishap" can be construed as the direct result of being *unaware* of one's biases in the moment of overlooking. Lack of awareness is a central feature of implicit bias (Kelly and Roedder 2008; Kahneman 2011; Banaji and Greenwald 2016). On dual-system theory, heuristic associations (cognitive shortcuts) of which we are unaware cause *stereotyping*, where System 1 hijacks System 2 (Kahneman, 24-25). Concerning race, implicit bias dredges up stereotypes – e.g., automatically associating "Black person" with "violent" – which implicitly influence judgments, actions, and attitudes (Kelly and Roedder 2008). Subjects are generally unaware of their biases and are much less likely to recognize their actions in the moment, much like what occurs when one takes an Implicit Association Test (Banaji and Greenwald 2016, 34-56). Thus, it seems plausible that overlooking is an example of IB without the empirical-psychological vocabulary.

But if it were true that subjects generally lack an awareness of their biases *in the moment*, then it would seem impossible that overlookers may become

aware of their biases. Is it plausible that Kim, for example, would really *never* realize or be aware of her possibly or actually biased behavior? This does not square with Kim's realizing what she almost did "all in a flash" whistling the start of "Swing Low, Sweet Chariot," where that means something to the effect that she instantaneously recognizes that she would have made a racially insensitive association between helping a black colleague with some labor along with whistling a slave spiritual. Were we to call Kim's association a bias, then it is unlikely that she recognized her action in the moment. Moreover, the sort of mishap we see in the case of Kim does not seem readily explainable by appealing to her hidden biases. We would have to say that she associated, somehow, "Black colleague + labor" with "the slave spiritual, 'Swing Low, Sweet Chariot'". But this would betray an unconvincing reading of bias as mapped onto what is a more sophisticated internal psychological conflict between committed-to beliefs and values (antiracism) and those which are reflectively eschewed (racist symbolism), even if both can be construed as accidental in nature. The point is that even if we were to construe Kim's behavior as a matter of IB, the suggestion that she may never become aware of her biases in the moment or even *ex post facto* does not map onto her lack of pervasively biased behavior.

3. Three Impacts of Overlooking

This section draws out three broad implications of the overlooking phenomenon. First, I raise a brief point about describing racism, where I sketch germinal ideas about future directions for thinking about racism which come to

light via overlooking. Second, I fill in some details about the overlooker's capacity to *internally regulate* her overlooking acts by offering a new and different description of amelioration.¹⁷ Third, I raise some worries about moral evasion implied by AI and IB and explain why no such evasion is available for the overlooker.

3.1. Description

The fundamental point I want to draw attention to is that overlooking opens our eyes to the way in which social-structural processes *play out* in individual psychology. Further, what overlooking shows is that there is no reason to plant a flag in either a social-structural or an individualistic approach. By “play out” I mean the effect social-structural processes have from *within* a subject's cognition. Now, this is a point about method – how racist phenomena are approached – and not a metaphysical point that social-structures only exist because of individual psychology. IB opens up a frontier to be explored, one which goes beyond either the social-structural or individual and locates the inter-relation between the two within a subject's mind; that is, the new frontier to thinking through racism is to see that the social *is* the mind, and the mind the social.

My point here is motivated by the following question: What more is there to do from a social-structural perspective that has not already been done by AI?¹⁸

¹⁷ Different to, that is, the descriptions of amelioration we confronted in chapter 2.

¹⁸ For a recent move towards individualism while still appreciating social-structural/institutional forces see Madva (2016).

We know this phenomenon exists, and we know social-structural forces can affect cognition.¹⁹ The idea is that we can agree with this view and expand out of it by focusing on how social-structural processes play out in people's minds. The next frontier I am pushing for has to do with a view like IB as the next evolution of social-structural analysis. Rather than *just* pointing to social structures, we consider their deepest, most insidious impacts on the mind, something we might not otherwise notice.

But why should we think through this more individualistic sense of racism when IB already exists, and this is the sort of thing that might make people racist beyond intentional hatred? The idea is that a more individualistic analysis is still consistent with a social-structural analysis and can be thought of as an extension of structuralism, even if it does not reduce to it. Social-structural forces are the *reason* why inexplicit racist phenomena like overlooking are so psychologically insidious. The point of talking about individual psychology in a new way is not just to pair together *structures + minds*, but to take it further and locate what exactly is going on in an individual's mind where structural forces play out. The benefit of doing so is not merely a matter of bridging individuals and social structures, but rather that the nexus between the two allows us to identify novel racist phenomena, phenomena with serious implications for those who might non-consciously think their academic knowledge and commitments put them beyond the reach of the structural forces of racism.

¹⁹ This is not to suggest the notion that racism is not captured by social-structures, but instead, that not all aspects are explainable by social-structures.

But how is it that social structures play out in individual psychology, how is such a thing possible? This is, at first blush, an odd claim. But recall from chapter 2 that we have in our possession a tool to show how this is possible, namely the *affiliation diagnostic system* (ADS). So, let's apply it here to the overlooking phenomenon. We know already that overlooking is a case of internal conflict. This conflict is mappable by the ADS, as that between someone's involuntary and voluntary affiliations. Unlike the cases described in chapter 2, there is no consistency or un-specificity concerning affiliations for the overlooker. Here we can see that overlookers disavow (even if they cannot completely escape) their involuntary affiliation to whiteness and consciously avow voluntary affiliations to out-groups and groups which acknowledge and take as their aim issues of racial justice and awareness. In virtue of the privileged norms of her birth, the overlooker is positioned to be racist and to incorporate that standpoint in accidental ways. But, although she is likely to some extent psychologically determined (her whiteness is inescapable), her choices of affiliation are not, having chosen to voluntarily affiliate with groups and sets of beliefs and values opposed to those forces. In thinking through this internal conflict via the affiliation diagnostic, we have a clear picture of what makes it possible that social-structural forces play out in individual psychology, namely by diametrically opposed affiliations.

3.2. *Internal Regulation and Amelioration*

Recall that unlike implicitly biased subjects, overlookers can regulate (with autonomous System 2 control) their unintentional racism and are not pervasively biased. But unlike IB, where someone else points it out and helps to train it out, here the overlooker does it all by herself. So, we might wonder: if overlooking involves unintentional/nonconscious actions, how will the overlooker know how to lessen it? I am going to demonstrate how, by an emotional response and introspection, overlookers can regulate their actions to show that there is a new sort of amelioration to consider, another to add to the running list from chapter 2. Regulation must come from somewhere; something must make it happen, but because overlookers are not pervasively biased, IB and AI are not going to cut it concerning regulation.

So, let's consider the psychology, a step-by-step story of how the overlooker might engage in such regulation. This is just a plausible story to reveal another sort of amelioration not considered, to my knowledge, anywhere else. Let's call the following description *introspective amelioration*, which means lessening overlooking by recognizing and regulating it. First, something might feel "off" in the overlooker, an inkling that something is wrong, having committed a racist act, but without knowing why she feels bad. She may have, say, non-consciously registered a victim's reaction. Perhaps she feels some

shame, even if she does not yet know its cause.²⁰ But if she is not aware of her action (and her failure to regulate it), then why would she feel bad? Simply put, we do not always have to be aware to have an emotional response.²¹ In any event, she does not yet know what she did that is the cause of her shame. The reasons for her unease are still nonconscious, just a vague sense that something feels off. But if she is feeling off, then it is realistic to suppose that she is motivated to figure out what is going on, because something is not squaring with who she takes herself to be. Although there is no *guarantee* that she will be motivated to unpack *why* she is feeling off, she is positioned to do so.²²

Given her unease, she may try to remember what she did – the overlooking action. Some part of her clearly registers her unease after the action. So, she takes a memorial inventory and traces back her uneasiness to the overlooking action: “I *did* do something wrong.” Now that she is aware of what she did, she is enormously motivated to understand what her motives might have been, but she is so convinced that “That’s not *me*” that she cannot imagine any valid motivates for her overlooking action. But her unease *will not* go away, so she may begin to think that some part of her *is* racist such that thinking “That’s not *me*” no longer makes sense. At this point she is really confused. She

²⁰ I will not detail the specific nature of shame here, but I have in mind something continuous with its being caused by a loss of self-respect after failing to live up to one’s own commitments (Rawls 1971, §67), and thinking less of oneself or a sense of guilt (Williams 1993, 93, 219-223).

²¹ If, *à la* Sullivan (2006), we can have unconscious habits, then why not nonconscious emotions? If we grant the former, we can grant the latter.

²² Not all cases of overlooking acts, nor even necessarily the majority, are accessible to the overlooker, but it is implausible that it is *never* accessible. All I need is one case where she can register her act and the rest follows.

does not recognize herself as the person who did the overlooking action. She may think, “Maybe some part of me *is* racist.” “I don’t want to be *that* person.” It dawns on her that something about herself she thought was true is not actually true, and that she should be on the lookout for this sort of thing.²³ Motivated by the force of shame and by a greater sense of taking care of themselves and others, overlookers will likely take personal responsibility for their unintentional racism and lessen its frequency.

3.3. *Moral Evasions*

Finally, where we propose ideas such as *in virtue of birth* or *in virtue of totally uncontrollable nonconscious thought processing* as the reasons why some people are racist, it seems to imply that we cannot hold them to account for their racism. “It’s those social forces.” “I was ambushed.” “It’s my brain automatically going about its business.” “I can’t help it.” These are all examples of *moral evasion* by which I mean *escaping moral accountability by appealing to exculpatory conditions*. The point is that subjects captured by AI and IB may evade responsibility. Whether by ignorance, deflecting blame to social-structural processes, or noting a lack of control, subjects may have (or may take) a “way out”. This is a consequence we may rightly be unwilling to accept.

AI models understand *active* ignorance as the original sin of racist actions taken as a result of (white) ignorance. It is not that subjects know better and did

²³ Although she is the kind of person who is likely to introspect on this, there is a *possibility* of denial here, that she wants to hold onto her self-image, and thus perhaps some motivation not to look inward at herself.

otherwise, an action for which they would be straightforwardly personally culpable, but that they (to some extent) *chose* to be ignorant in the first place – i.e., they are *actively* ignorant – for which they are culpable for all actions taken out of ignorance. Yet the connection between the socially structured ignorance and the *personal responsibility* of the ignorant individual – i.e., taking responsibility for oneself and/or holding oneself to account – seems tenuous, in that we might suppose that individuals could appeal to being caught up in vast forces beyond their personal control.²⁴ Conversely, we seem to be unable to hold to account those who really do know better and still do racist things, such as the overlooker, because they are merely lumped in with a mass of white ignorance. It is not that individuals cannot be held to account for their participation in harmful structures *at all*, but that faceless structures seem to both lessen personal motivations to change (no matter what they do, they are still part of the structure of whiteness) and taking responsibility (they are at fault in virtue of the structure rather than their own actions).

Further, given that we plausibly assume that control is a necessary condition for moral responsibility, then on models of IB we have another tenuous situation, as a lack of conscious control is a feature of those views. By appealing to not being in complete control of one's actions and behaviors, a subject may

²⁴ Whites, then, could be said to be instruments of the “structural will” in the sense that corporate agents are instruments of a “corporate will” (as in French 1995). Where, in virtue of her whiteness, someone follows the structural will, she acts within a structurally oppressive or unjust system. The implication is that under these conditions subjects are not themselves responsible when the supervening “script” goes, or is, wrong. They can effectively relinquish moral responsibility, given that the structure sets into motion determinants of individuals’ agency.

claim this as an excuse for evading responsibility. Subjects, then, may blame their automatic biases or an insufficient IB training seminar, divorcing themselves from any personal role or agency – “It’s not *me*, it’s my brain.”²⁵ They are unable to endorse their actions with any reasons *à la* System 2 regulation, and so those actions fall outside their conscious control. So, given that implicit bias short-circuits a subject’s control, she can evade moral responsibility for that action.

But notice that overlookers have no recourse to the foregoing evasions. Most obviously, they cannot claim to be ignorant (active or otherwise). They are not simply informed to some degree, but they knowingly (albeit unintentionally) cause, or nearly cause, racist events. Neither can overlookers deflect blame away from personal responsibility to social-structural forces, because (from 3.1.) these forces operate from within in the mind. Finally, that the overlooker fails to enact the knowledge and care she already has entails her capacity to do otherwise. Overlooking includes both the capacity to do some action φ or do some other action ψ instead, as well as the capacity to reflectively endorse some set of reasons for some action. To fail in the enactment of knowledge reveals in the overlooker that she has a chance to regulate her action in such a way as to align with her value scheme. Given that overlookers can or would have done otherwise, it cannot be claimed that they lack conscious control over their racist actions.

²⁵ There is a growing literature on the issue of IB and moral responsibility. For an anthologized treatment of the issue see Brownstein and Saul (2016) (especially Part 1). Unpacking this issue is beyond the scope of this chapter; I only wish to put pressure on what IB *implies* with regard to responsibility.

The ultimate point is that phenomena like overlooking show us that personal responsibility is something to be on the lookout for, and that morally evasive implications and tactics ought themselves be evaded. No matter how much we know and care, we are not beyond racism, just like anyone else who does *not* know and does *not* care (or those who are post-IB training who might think they are done – “It’s *those* white people over there, not me”). This is true even in a more personal way than what structuralists might suggest – that one is just racist because of one’s participation in systems of racial oppression and injustice. This latter point has a damaging effect on taking personal responsibility for oneself, precisely because there seems to be no hope, and thus no motivation, of ever escaping the system.

In that vein, the question overlooking raises as concerns responsibility is: how does or how can one take *personal* responsibility for social-structural ills? Thus, overlooking does not resolve any moral issues, but instead shows us something about responsibility – that responsibility, like overlooking racism, is neither *just* personal nor *just* social. If it is not just the individual responsible for herself or the groups and structures in and from which she participates and benefits, then where does responsibility live, where is it located? These questions are answered in chapter 4, where we will continue walking a tightrope – as we have done in this and the previous chapter – and show that we should not privilege the personal over the social or *vice versa*, but instead re-frame the idea

of responsibility itself in such a way that avoids thinking exclusively in terms of either the personal or the social.

Conclusion

Let's speak plainly and practically. Is it psychologically realistic to suppose that self-shaming academics (and the "woke" crowd in general) walk through life in a *constant state of shame* based on their oppressive racial status? What seems more likely is that constant vigilance and utter self-shame are psychologically unsustainable outlooks and states of mind. Phenomena like white fragility, bad faith wokeness, and virtue-signaling all depend on not just recognizing but *psychologically living in* a state of original sin. No one can maintain that frame of mind at all times, and those who might maintain they do are the very ones susceptible to overlooking their own racism. So, what we end up with are folks who likely think of themselves (consciously or not) as not *those* (racist) white people, a comforting if misguided sentiment they take to be a truism of themselves. This implication is one we might rightly count dangerous.

One does not have to be Kim, that is, one need not be an expert on race and racism. One can be minimally knowledgeable and minimally care. But the problem is that if one buys into the sorts of assumptions that AI and IB lay out, then one might end up a member of the bad faith woke virtue-singling crowd. This creates, it seems to me, an artificial rift between the "in the know" and the "not in the know". The *in the know* white folks might just think they are not the ones who fall within the scope of AI or IB; rather, it is the white folks who are *not*

in the know. But is this how things *really* stand? The practical worry is that this promotes “but not me” sort of thinking, that is, the sort of thinking that one has arrived – that one knows better and cares, which is exactly the point overlooking addresses. But also, in important ways, this takes ideas of racism down from academia’s ivory towers precisely because no amount of theoretical learning is sufficient to prevent racism in a *real* and *personal* way, not solely in virtue of one’s abstract *whiteness*. It is very human to draw a divide between knowledgeable whites who are devoted and “know better” *versus* the masses who do not know better. Overlooking shows that this artificial division is split, the barrier between academia and “everyone else” dissolved. Those who care and know better sometimes occupy “the other side” of that divided line. *You* may have spent some time on the other side. What is at stake here is just a matter of being accountable as a fallible human, no different in kind than other humans who do not have special theoretical knowledge. If people are to maintain unrealistic outlooks, then proliferations of bias, ambush, fragility, and overlooking will ensue; we may give up, because it seems so hopeless.

But in tandem with the utter impracticalities of the always-vigilant self-shamer approach, we arrive at a dangerous blindness, because no one has that always-vigilant ability, and where we fail, the implication is that we are not at fault. And if we cannot do this without impossible standards, then we get more and more instances of the overlooking phenomenon. The first step towards amelioration, and the performative function this chapter attempts to execute, is

awareness of the overlooking phenomenon. No one in good faith can assume “I’m all through with racism, because I know better; I care.” The ultimate point of this chapter on overlooking is to care about it, to see it. Be on the lookout. Be vigilant. But not so as to be perfect, but instead, in admitting imperfection, retaining better control in one’s actions, and changing who one is. If we cannot take the first step towards awareness, what results is likely to be an un-ameliorated proliferation of overlooking and overlookers, not to mention actively ignorant and implicitly biased persons. These are people you know. They might be you; they might be me.

Chapter 4 A Continuum View of Responsibility

Introduction

Philosophers of race often construe racism as a moral concept and apply the conceptual vocabulary of ethics to the problem of racism. This chapter attempts something different by reversing that order. It asks: *what if we took issues that arise in racism and see whether they can help us think more generally about ethics, and its conceptual arsenal, as such?*¹ It answers in the affirmative and focuses on responsibility. Thus, this chapter is a meta-ethical inquiry into the idea of responsibility. What constitutes responsibility? What are its proper targets? Where is it located?

The previous chapter presented excuses called *moral evasions* – escaping accountability by appealing to exculpatory conditions. We saw that there might be reasons that actively ignorant and implicitly biased people can evade responsibility. There is privilege in the ability to do so. In effect, these people can dictate the extent to which they are responsible. It would seem, then, that the ability to evade is made possible by having the weight of privilege on one's side. There is real power involved here. But what is the relationship between power and responsibility more generally?

¹ This is different to understanding responsibility *in terms of racism*, in the same way we might understand something in terms of gender roles in feminist philosophy or in terms of economic relations in Marxist theory. Of course, we can use a critical methodology that interprets responsibilities in light of unjust social structures and racism. But that is not what I will do.

In all the literature on responsibility there is very little in the way of a substantive view of power, aside from narrow senses of (metaphysical) “power” as an agent’s ability to cause actions.² What many responsibility theorists instead analyze are the various nuances of individual and social responsibility. What generally ensues is running in metaphysical circles, oscillating attempts at offering the “right” view of agency meant to resolve whether individuals, groups, or societies are the appropriate targets of responsibility for some harm(s). But what if we weren’t forced to choose between these targets?

In chapter 3 we saw that social forces play out in individuals, that the social and the individual are connected in the *overlooker’s* mind. Here we consider how power relations instantiate in groups composed of individuals. Thus, whereas for the overlooker the nexus of the individual and the social is the mind, here *power* is the nexus. So, this chapter will show that re-framing responsibility by reference to power avoids the metaphysical oscillations noted above—*power does not force us to choose*.

But beyond avoiding the complications of the individual and the social, focusing on power has real practical purchase. We should want to go beyond saying we’re only responsible just for the things we do. By re-framing responsibility in terms of power, we can say that we’re responsible for *more* than *just* what we do, but also that some are *much more* responsible than others. This

² To my knowledge only Young (2006), McKenna (2018), Mackenzie (2018), and Oshana (2018) focus on non-narrow, non-metaphysical senses of power. All generally have in common the idea that responsibility is contoured by power dynamics that emerge from unjust structures of racial, gender, and/or ethnic oppression. To reiterate, that is not my approach here. (See note 1.)

might get *people* to actually *take* responsibility. So, I offer a notion of responsibility that is tethered to the ground, that has practical consequences. It is a notion that acknowledges that while not all may enjoy equal power in *all* aspects of life, that shouldn't disincentivize them to reject the notion that they bear *some* responsibility.

1. Surveying Options

This section critically appraises three general positions on responsibility in the literature: *individual*, *collective*, and *shared*. It highlights virtues of each position but also presents what I am calling "complications" each runs up against. These complications are the result of how each position metaphysically frames agency, which comes in tow with theoretical and practical problems we may rightly want to avoid. I argue that there are good reasons to re-frame the conversation about responsibility to something beyond the focus on individuals, collectives, and aggregated members of groups. This motivates locating responsibility on a continuum, a view which avoids these consequences, and adds some nuance of its own to the conversation.

1.1. *The Individual*

Let's look at individualist views. We will look through both a metaphysical lens (what it means to be a responsible "agent") and a moral lens (what it means to be held accountable). Both senses, of course, look at individual persons. The focus on individual persons is what I am portraying as common

amongst these views, as I cannot exhaustively present the variety of existing positions and all their nuances.

Identity, Action, Morality

Let's start with the metaphysical lens first. To be responsible is a question about the composition of a moral agent, that which makes her a viable candidate for responsible agency: e.g., that she has a will, that she can author her own intentions and actions, and so on.³ This is about the identity of an agent *qua* agent; one is responsible in virtue of her action's being attributable to her. She is the "author" of that action, where "authorship" is intentionality: her motivational desires and her freedom to do otherwise. So, she is responsible where her actions and behaviors are directed of her own volition (Frankfurt 1971; Wolf 1987, 1990; Watson 1996, 234; Fischer and Ravizza 1998; Smith 2005, 251-252; 2008, 370; Sripada 2016, 1211, 1216). That she is the author of her actions is what makes her a *responsible* agent.

Now that we know why an individual can be responsible *at all*, we can speak to why she can be *held* responsible. Here the focus on individual responsibility centers on what responsibility *for* something means morally and/or politically – e.g., violating some duty to act ("prospective" responsibility) or having violated some duty for an action that already happened ("retrospective" responsibility).⁴ Though individualistically centered, the

³ To be clear, the question here is not about whether someone is responsible because it is good *for her* (an egoism of some sort), but rather, whether she stands as someone who could be responsible *at all*.

⁴ I am drawing this "prospective/retrospective" distinction from Zimmerman (2015).

position acknowledges someone's membership in a moral/political community. One's actions sometimes affect others. Someone can be held responsible for her actions based on, for example, others' "reactive attitudes" (Strawson 1974), our failure to "comply" with the duties of our "social setting" (Watson 1996, 229), a violation of contractarian duties (Scanlon 1998), or our actions' negative outcomes (Smart 1961, 1973). So, if the moral community of which I am a member values promise-keeping and thinks it a duty to keep promises, then my breaking a promise is a moral violation. Thus, others can hold me responsible.

A Virtue, a Complication, an Implication

A virtue of these views is that they are explanatorily compelling. What would it mean to act without assuming a will, desires, or beliefs? How could we assign moral status to actions without these? Where *S* did action *A*, it came from *S*'s will, and so *S* is the one who is held to account and no one else. But if *S* did not do *A*, if she did not will *A* to be so, then *S* is not at fault.⁵ Meeting certain duties not to harm others is useful for determining whether an agent is responsible for what has been done or what she failed to do. Unless we want to do without ever holding persons to account, some basic notion of *how* and *why* persons are accountable is indispensable.

But while features of personal responsibility are indispensable, the position runs into complications where structural harms and injustices are concerned. Call this the *structural complication*, by which I mean structural harms

⁵ Obvious hedges notwithstanding—e.g., genuine actions of omission.

likely cannot be tied to individual actors. *I* did not author, intend or directly cause a structural injustice or its harmful outcomes, nor did *I* set these social processes into motion – e.g., world poverty, environmental harm, educational inequalities, unequal access to healthcare, and so on. It does not seem likely that we can link those *structural* harms to specific, singular individuals. If we hold that these are morally significant states of affairs, yet no one individual can be blamed for them, then individualism fails to offer a sufficient notion of responsibility.

Individualism, further, runs an implied risk that individuals may simply *deny* or *ignore* their potentially complicit roles in maintaining structural harms, in that one can plausibly deny one's intentional role in perpetuating structural harms. This lets individuals off the moral hook. So, something more is necessary beyond a view of individual moral agency.

1.2. The Social

Perhaps the structural complication and its implication are avoidable by focusing instead on groups. If *one* of us isn't responsible, maybe *we all* are. So, let's turn to positions on social responsibility: *collective* and *shared* responsibility. I use "social responsibility" to capture these positions, because each depends on there being at least two (or more) individuals. The point of contention between the two is whether or not responsibility is distributed to all members of the group, where "distribution" means doling out responsibility to individuals who

populate the group. Collective responsibility does not always assume so and shared responsibility does. Let's turn to collective responsibility first.

Collective Responsibility and (Usually) Non-Distribution

Genocide, environmental harm, unequal educational opportunities, disproportionate access to healthcare, and so on, are not outcomes that are neatly ascribable to individual agents. But we might rightly want to ascribe responsibility to *something* for these issues. Here is where collective responsibility helps. As before, I will not encyclopedically overview collective responsibility. Instead, I just want to point to what its views have in common: a holistic, non-individualistic approach. The idea is that unified wholes ("collectives") can be held responsible (usually) *without distributing responsibility* to individual members of that unified whole.

For example, it is the car company that is responsible for an environmental harm, and not necessarily all its individual corporate agents. Or it is the nation that is responsible for genocidal acts, unequal educational opportunities, or disproportionate access to healthcare. The reason is that collectives are cohesive, organized groups which undertake collective actions, where the group itself is an "agent" distinct from its members (French 1979, 1984, Chs. 3 and 4, 1998, 37; Corlett 2001, 575; Mathiesen 2006; List and Pettit 2011). Consider Peter French's (1984) view of corporate agency: corporations are moral agents because they act intentionally, have the capacity for rational decision-making, and can respond to reasons (ibid, Ch. 3). A group, whether a company

or a nation, has a decision structure that can be enacted *despite* its individual members. If a group, as a whole, can respond to reasons and collectively act, it is not so counterintuitive to suppose it can be held responsible as a whole.

But here is where the question of distribution arrives, causing things to get murky. While some theorists think collective responsibility is not dissolvable to individual actors of a collective (Arendt 1987) or unanalyzable in terms of individual responsibility (Issacs 2006, 2014), others find room for responsibility of *at least a few individual* actors of a collective (French 1998, 25; Mathiesen 2006; Mellema 2006). So, collective responsibility is generally *non-distributive*, but *not always* non-distributive. A corporation may be responsible for some past harm at *time₁*, but perhaps no *current* members are individually responsible because its employees have been entirely replaced at *time_{n+1}*. But where a nation commits genocide, we might rightly want to say that *some* individuals (perhaps some leaders) are at fault, while others may be absolved of responsibility.

A Virtue, a Complication, an Implication

Two virtues of collective responsibility are its *explanatory power* and its *praxis*. The former can show why a complex phenomenon can be unified or reduced by understanding it in terms of something simpler. So, collective responsibility is powerful in this sense because we do not have to search through the messiness of individual intentions or agents. Instead, we cut through the details to hit the basic principle (membership in the collective), that which all the individuals have in common, by which we then assign responsibility. But

collective responsibility is also practical: it allows a principled basis on which we can argue that whole groups should, for example, make good on social programs such as reparations and affirmative action, things we might find practically desirable to do.

Yet, collective responsibility runs up against a *distributive complication*. The position slides between distributing responsibility to individuals and not. There is something of a dilemma here. On the one hand, where we don't distribute responsibility to individuals, we run the risk of absolving them, good and bad, of responsibility whether they directly participated in the collective's harms or not. While this is sometimes a virtue of the position (completely innocent persons are faultless), it lets off the hook those who may have directly contributed to the collective's harms. On the other hand, where one *does* distribute responsibility to certain members of the group, we might for example both maintain the nation's responsibility for genocide and distribute *some* responsibility to the organizers and killers. But where does the distribution stop? Is it enough to single out some token actors from the collective whole?⁶ But if we name individuals, we lose explanatory power, the virtue of not having to sort through the details of every case, the broader context, and the intentions of individuals. These are the sorts of things we were not supposed to have to worry about in the first place.

⁶ It may be worth wondering, also, whether the view reduces to personal responsibility if we slide all the way down the distributive slippery slope. It also seems likely that drawing a line in the sand would raise difficulties about borderline cases: who is "just enough" responsible and who isn't?

A further complication is a practical *motivational disincentive*. If all are absolved or (if we do not know where to draw the distributive line) not all are absolved, then many (all?) will not feel personally responsible. And feelings matter, because then individuals will not *take* responsibility – i.e., they will not change their harmful ways. Individuals might think they do not “own” any responsibility – it’s owned by the collective. The problem is reminiscent of moral evasiveness (as discussed vis-à-vis *active ignorance* in chapter 3), because individuals have a ground from which they can point their fingers to society instead of themselves, effectively turning away from the harms their society causes. Whether it is genocide, a car company, or the structural problems of a nation, it is not farfetched to suppose individuals will look anywhere but to themselves. Just as in chapter 2 where we noted the role of the philosopher as a functionary of humanity, and that public reaction matters vis-à-vis one’s theoretical and practical goals, here we have a similar issue. When we have a notion of responsibility that is so disconnected from individuals’ personal responsibility, no one will feel any direct ownership of the collective’s actions. Why care when it is the car company or nation that is at fault?

Shared Responsibility and Distribution

If we want to avoid the foregoing complications, yet retain some emphasis on groups, then we will have to look elsewhere. Notions of shared responsibility connect responsibility to individual members constitutive of a group, which (unlike collective notions) is not necessarily a cohesive or organized group that

acted together to cause some harmful event.⁷ So, whereas collective responsibility is non-individualistic, shared responsibility is individualistic.⁸ And whereas collective responsibility is generally non-distributive, shared responsibility is *always distributive*.⁹

As before, consider the general features of the view: a methodological individualism. Moral agency is attributable only to individuals in a group, not the group *qua* group. So, where *A* is some group action, *A* is only a group action insofar as it was undertaken together by individuals in that group. Each member of the group is, then, responsible for *A*. Think of a conspiracy to commit murder. It is not just the individual who did the murderous act who is responsible, but rather that the responsibility is “distributed” to all those who participated in that group. Whereas collective responsibility is not generally connected to individual members of a group, shared responsibility is. So perhaps shared responsibility can recapture that lost explanatory appeal of the personal without its anti-structural drawbacks, and retain the social focus of the collective without its impractical results.

On *shared agency*, groups “have” agency because their individual members share the same sorts of intentions or attitudes and act together accordingly, not *despite* their individual members’ intentions or attitudes. Shared agency is

⁷ See May (1992, 106-107).

⁸ Though it connects responsibility to a group’s individuals, it is not synonymous with individual responsibility (as presented above) because, as noted, it depends on the existence of at least two or more individuals whereas individual responsibility need only depend on one.

⁹ See Feinberg (1968, 647), May (1992, 37-38), and Issacs (2006, 61; 2014, 43).

captured in a few different ways. Sometimes it is captured by a cultural “climate of attitudes” motivating persons to act (May 1992, 46, 53-54). There is certainly such a climate in the conspiracy case, that each contributes towards a positive attitude towards the murder. It could also be construed in terms of shared intentions, that where a and b intend to do A , a and b are responsible for A . (Bratman 1999, 121; 2014, 103). There was a shared intention to murder because the parties all intend the same action. Or it can be seen as dependent on being “socialized as a member” of a group (Silver 2002, 299-301). This means that whether we like it or not “we cannot choose to stop viewing ourselves as members” of a group (294). Each member of the group, having been socialized as a member, shares in that group’s murderous identity, and has ownership over its actions. Identity as “shared” by group members may also ground shared “duties to respond” (Radzik 2001, 466). In our toy example, this may involve the normative peer pressure to *not* back out or to *not* prevent the murder from happening.

Suppose we take May’s (1992) “climate of attitudes” approach. Attitudes cause a “climate” that can make it such that certain kinds of harm are likely to occur in some community (46). For instance, a climate of COVID-19 vaccination wariness may result in the deaths of others (and/or oneself). Each individual who shares the intention not to vaccinate shares also the harmful results of abstention. Such a “climate of attitudes” *motivates* shared action, where each

member at least partially plays a role in contributing to the wrongs resulting from the action.

If shared agency is granted, then members of a group act together (directly or indirectly) to bring about some harm or fail to prevent it. Thus, each of those members shares in the responsibility for that harm (Zimmerman 1985; May 1992, Ch. 8; Sadler 2006; Young 2004, 2006, 2011; Darby and Branscombe 2014). That is, unlike collective views, here responsibility is *distributed* to individuals in the group. Though some think responsibility should not be distributed in equal measure (May 1992, 10-11, 38-42; Young 2006, 125; 2004, 381, 383-388), others argue that it should (Zimmerman 1985, 115). Despite this disagreement, the overarching point is that all agree that each member of a group is responsible for the outcomes of a harmful shared action, whether they contributed directly or indirectly to those outcomes. So, while not all of the conspirators participated in the physical murderous act, each bears responsibility for the harmful outcome, having played a role in planning the act. Similarly, members of the anti-vaccination community each bear responsibility for the harmful outcomes of abstaining from getting vaccinated.

A Virtue, a Complication, an Implication

A virtue of shared responsibility is that it is *integrative*, by which I mean it captures why we have reason to think individuals, interconnectedly constitutive of groups, are each responsible for harms caused by the group. No one is let off the hook for group wrongs, which avoids the implied denial of collective

responsibility. And the position provides a useful tool for unpacking the nuances of broader cases like structural harms. Suppose members of a governing body pass legislation whose outcome is disproportionate access to healthcare for some citizens. It seems clear enough that those members each share responsibility for this outcome. But what of those who voted for those lawmakers? Here the participation in passing legislation is not overt, but it is still contributory, because those lawmakers would not have their positions were it not for their voters. Shared responsibility provides a groundwork by which to distribute accountability to each person involved in the harm regardless of whether their role is direct.

But although the view lets no one off the hook, it faces a *participation complication*, by which I mean two things. First, shared agency, on which the position depends, cannot account for what we can call *epistemically isolated cases*, that is, cases where agents know nothing of committed harms, which may be cases of genuine ignorance. Second, shared agency cannot account for those who *cannot* participate in taking up responsibility to prevent harms. This in turn implies an *excessive rigor*, that distributing responsibility to each member of a group may be unfairly burdensome.

For instance, consider an isolated Hutu Rwandan farmer. He tends to his goats, plays no role in the Hutu regime, and knows nothing of the harmful events it caused. On the foregoing view, although the farmer is not involved in any Hutu genocide, he still shares (partially) in the responsibility for those

harms – perhaps not because he did nothing *at all*, but because he did nothing to *prevent* any of those harms. This seems to offend against fairness, because he does not contribute in any relevant sense to the harms committed by the Hutu regime.

But there is a second sense where participation is likely not possible *at all*. If global issues such as environmental harm are something for which we ought to together share responsibility, then it would seem that taking up the torch might be available to only a select sub-set of a population – i.e., participating is not a live option for everyone. Those who cannot afford an electric car or must buy plastic bottled water due to unpotable water in their community, cannot share in responsibility for preventing environmental damage. Are they at fault? The point is that shared responsibility seems to imply unfair and, in some cases, tone-deaf burdens.

2. Responsibility on a Continuum

This section aims to avoid the foregoing consequences by locating responsibility on a continuum of power. Very simply, the more power we have, the greater our responsibility. I arrive at this view by thinking through a series of continuums, and what they imply, relating to harm, power, and agency. Along the way I will show the superiority of this view by showing how it avoids the imbroglios of the literature's camps. Now, to be clear, I am not going to offer any hard-and-fast rule why or that someone or something is responsible; no normative theory of responsibility will be offered below, and I'll say nothing of

moral blame. Rather, I am considering a different way to think about responsibility, a different way to frame an ethically relevant word.

2.1. *Continuum I: Harm*

Let's begin by distinguishing two senses of responsibility. Suppose I'm waxing cars at a car wash. I'm *causally* responsible for the wax job because my actions caused it. Now, I didn't have much choice in the matter if I was told by middle management to use a particular car wax brand. So, though I'm causally responsible for the wax job, I didn't have much of any choice to use that particular brand.¹⁰ Now suppose the car wax is defective and ruins the paintjob of the cars to which it is applied. I've caused a harmful effect, not in the odd sense that the paintjobs are "harmed", but that the *customers* are harmed.

But what is a harm?¹¹ While I will not offer any novel definition of "harm", I take loose inspiration from Joel Feinberg's (1984, Ch. 2) nonnormative notion that harm is a "setback to interests."¹² Let's unpack this. A setback to interests is a consequence of *something*—i.e., it is an effect. But this does not happen in a void—it happens *to* someone or some group. And what "happens"

¹⁰ The question of "degrees" of agency will be elaborated in the next section. Here I'm only attempting to begin with a plausible case that degrees of harm matter where responsibility is concerned.

¹¹ The reader will find in a moment that I borrow from Feinberg (1984). But my thinking is continuous with Bradley's (2012) desideratum that an account of harm should not assume that harm is always morally wrong (395). Following Hanser (2008, 421), I think that harm is morally significant, though I do not begin there.

¹² Harm as a setback to interests relies, for Feinberg, on a so-called "counterfactual" account of harm, according to which where *S* is harmed, *S* is in a position where her interests are set back in a way they otherwise would not have been. See Feinberg (1986). See also Thomson (2011, 447-450) for another version of a counterfactual view. I will not spend time here defending that account. See Klocksien (2012) for a defense.

is that their *autonomy* is negatively affected. For now, “autonomy” just means the capacity to make an informed choice to act on an option from some set of options without any undue external influence. Moving forward, harm will include a notion of autonomy, namely that it was negatively affected.

So, let’s say that a *harm* is, *prima facie*, some effect that is attributable to the causal capacity of an agent or agents, where that effect has potential negative consequences for individuals or groups. I do not mean at the moment anything explicitly moral about harm, though of course this is not to deny that harm has moral implications. It would seem that some setbacks are greater than others, that harm is delineable by degree. So, we might further develop ideas about degrees of harm involved on a continuum running along two axes: *severity* and *quantity*. The former has to do with *how* harmed someone or some group is. The latter has to do with how *many* are harmed. What I will attempt is to make a plausible case, by example, that harm exists on a continuum.

Let’s explore this further by marking off different degrees of harm by *severity*. How harmed are individuals or groups, that is, how much has their agency been affected? Consider the following examples.

GLUTTONOUS TODDLER. I prevent a toddler from eating three pounds of ice cream in one sitting.

VIOLENT CRIME. I commit premeditated murder against a completely innocent person.

DISCLOSURE. A company does not disclose information about their product, preventing consumers from making fully informed purchases.

SYSTEMIC. A region practices redlining, encourages gentrification, and sets *de facto* limits on equal access to healthcare, all of which disproportionately affect a sub-set of the population.

Though my preventing the gluttonous toddler from eating three pounds of ice cream in one sitting may be a setback in some minimal sense, it seems she is not *very* harmed (if at all).¹³ There must be some reasonable construal of autonomy where harms are concerned. It is not unreasonable for me to prevent the toddler from eating that much ice cream, because she lacks the autonomy to make an informed choice. But it *is* unreasonable of me to commit a physically violent crime against a completely innocent person because it is a clear-cut case of totally extinguishing someone's autonomy. The company's failure to disclose information about their product is another unreasonable denial of autonomy because consumers were prevented from making a fully informed choice about their purchase. Here the harm is less severe than being a victim of a violent crime, but more severe than the toddler example, because the former eliminates autonomy, and the latter reasonably limits it. The systemic setbacks unreasonably limit autonomy because they unfairly restrict opportunities, choices, and the basic needs and interests of a whole swath of individuals. Here is a harm still less severe than the total elimination of autonomy but perhaps

¹³ Some form of paternalism is likely justified in such a scenario.

more severe than the consumer example, as the range of setbacks is far broader than the “choice” of purchasing a singular product.

Let’s now explore different *quantitative* degrees of harm. How many are harmed? Before doing so, however, I want to make it clear that I am not suggesting that the greater number of persons who are harmed means that the harm is therefore more severe, that quantity always directly correlates with severity. As I am not offering an explicitly moral analysis of harm, the reader should not assume any consequentialist calculus, rather we are only after how harm is measurable along two dimensions.

Now, we have already seen an example of *one* person being harmed: the victim of a violent crime. But the other two examples broaden harm’s numerical scope. So, let’s consider some more examples to further substantiate the harm continuum in terms of quantity.

VEHICLE. Purchasers of a vehicle only sold in the U.S. are harmed because it was manufactured using an unreliable transmission that was selected by higher-ups because it was a cost-effective option.

SHOES. A shoe manufacturer employs sweatshop laborers to mass-produce its footwear.

LAW. In 1973 the United States Supreme Court decided in the case of *San Antonio Independent School District v. Rodriguez*¹⁴ that it is not unconstitutional to fund public schools based on local property taxes.¹⁵

In VEHICLE, the quantitative scope is restricted in at least two ways: to the U.S. and to only those who purchased the vehicle. This is a case of direct harm to a group of faceless consumers, smaller in scope to the examples of LAW and SHOES, but many are nevertheless harmed. SHOES is similar in scope to LAW, where great numbers of persons are harmed – i.e., sometimes harm is society-wide, baked, as it were, into social structures. In LAW, the result is that equal funding among public schools in the region was not equally distributed, leaving schools in poorer districts significantly underfunded and their students deprived and disadvantaged in comparison to schools in wealthier areas. It amounted to indirectly calcifying a structural harm, affecting many disproportionately.¹⁶

To my knowledge, all three positions on responsibility from section 1 take harm for granted as an assumed component of a theory of responsibility. Even where we hedge by saying that some may not be equally responsible for some harm, it is not at all clear *why*, against the backdrop of harm, this is so. Perhaps because the debates are so metaphysically-framed (what makes for a moral agent, what a collective is and how it is an agent, and how shared agency works)

¹⁴ 411 U.S. 1 (1973).

¹⁵ That is, it was decided that it was not a violation of the Fourteenth Amendment's equal protection clause.

¹⁶ I say "indirectly" because the Justices may not have "created" the structural harm, but rather, solidified it.

that they end up presuming some of the necessary essentials without much description. But harm is not an amorphous phenomenon to be simply assumed. If harm is taken for granted, not only do we miss out on its nuances, but we also run the risk of unfairly allocating it.

The importance of offering a harm continuum is precisely that it adds some sophistication to our understanding of responsibility. That we can appreciate that not all harms are the same, that some are more severe than others and some affect greater numbers than others, gives us greater purchase in knowing *how* responsible individuals or groups are. We might say we are more responsible for harms with greater severity, or that we bear a greater responsibility for preventing a large quantity of harm. But given that harm has been tied to agency, as an effect that we cause, we might rightly wonder whether degrees of harm correlate to degrees of agency, whether the capacity to cause greater or lesser harms has something to do with the “amount” of agency we have. This is a question to which we turn next.

2.2. Continuum II: Power and Agency

If the harm continuum is granted, then how can it be caused to greater or lesser extents? Let’s garner some insight from José Medina’s (2013) notion of responsible agency which emphasizes an oppressor/oppressed relationship. For Medina, mitigating systemic injustices requires that oppressors, or those with disproportionate power (i.e., privilege), know the ways they are socially positioned in relation to those injustices (133). Beyond oppressive systems of

racial or gender injustice, is there something more general to say about responsibility as such borne out by Medina's insight? The proportionality of power found in the oppressor/oppressed relationship helps us to think about responsibility in general – there is something special about our responsibilities given the sorts of persons we are. A continuum of power and proportionality has to do with how privileged status – i.e., how much power we have – corresponds to the effects we can propagate.

Having power over others can lead to harms as negative outcomes, so let's first consider why *the more power we have, the more agency we have*. By "power" I mean the greater or lesser extent to which we¹⁷ can change something or effect some result, results that have implications for other individuals.¹⁸ So, for instance, power can be as minimal as my preventing a toddler from eating three pounds of ice cream in one sitting, and as maximal as a CEO's decision to put that brand of sparkplug into the engines on the assembly line or a nation's decision to go to war. The point is that power comes in degrees based on its range of causality and effects. The nation going to war has a tremendous capacity

¹⁷ I use the first-person plural pronoun "we" here instead of "individual", "group", "collective", or "aggregate" not just for simplicity's sake, but also because I want the scope of power to be broad. So "we" could be a person, a group, a group within a group, a collective, a nation, and so on. Thus, we can speak of an individuals' power within a group, a group's power as a whole, or the power of sub-groups within a group (e.g., a board of directors, who within a larger group have more power than others in that group).

¹⁸ My notion of power, then, is different to "getting what one wants" (Goldman 1972, 222-223; Hobbes 1994, Pt. 1, Ch. 10). Nor does it have to do with doing what one wants in the face of resistance (Weber 2013). Because I will, below, claim that power is a property of agents, my notion is contrary to Arendt's (1972) who thinks that "power is never the property of an individual; it belongs to a group..." (143). And I do not quite mean by "power" what Foucault (1978, 1980) means – i.e., I do not mean it as operating in a matrix of truth, knowledge, and ethics. But my view is continuous with Russell's (1996) notion of power as the "production of intended effects" (23).

to cause effects, choosing to mobilize troops and affect the regions wherein they fight. The CEO's sparkplug decision affects all cars manufactured in the plant. The assembly line worker's power is much lower than that of the CEO's. She does not choose which sparkplugs to put into the engines; she only puts them in.

Now, the kind of power we have, and its degree, is only possible under certain conditions. The exercise of power is *contextual* to the group in which it is exercised. For instance, Supreme Court Justices have a lot of power in the context of jurisprudence, the CEO over a car company, the governing body of a nation over its citizens, and so on. Thus, it is in virtue of our *affiliation* with a specific group *that* we have the kind of power we have. People with power within those groups achieve what they achieve within contexts. Responsibility *qua* groups must, then, include some notion of the relative power of the individuals involved: a *continuum of power*.

Harms (effects) are caused by agents. Power is a property of agents. It depends on agency for its exercise. I mean by "agency" a stronger notion of *autonomy*. It is (1) *synchronic*: one's narrow ability to cause any single action or event. And it is (2) *diachronic*: an ability to cause a series of actions or events. Where agency is *synchronic*, it involves at some time slice my ability to cause an event in any one choice set, a range of choices one of which we *do*. Where agency is *diachronic*, it involves a series of choices made over time, instrumental steps each of which we have synchronic control over — choices we have *to* cause actions or events. One is an agent on this construal where one can act according

to her own principle in free self-determination.¹⁹ Negatively put, it is the capacity to cause our own actions without undue external influence.²⁰

Permit me to elaborate further. Agency is the ability to cause a range of effects in any one instance of a possible event scenario.²¹ And if I have a high degree of autonomy, I can narrowly choose, for example, to go to the health foods store to purchase groceries or to McDonald's. I choose one or the other and then cause the event to happen by doing one or the other. More widely, I can instantiate my values about a healthy lifestyle over a series of health-conscious choices where I have the money (autonomy) to go to a gym, shop at Whole Foods, see a nutritionist and so on. For any one choice, narrowly, I might have a greater range of options, and for a series of choices, widely, I can instantiate a greater range of options in my series of choices over time. An agent may have a greater or lesser degree of *autonomy*, which instantiates both choice features.

While both agency and power have to do with the capacity to generate an effect, power is *relational*—exerting influence over others²²—whereas agency implies control over *oneself*, and not necessarily others. That we are in control

¹⁹ Though my sense of autonomy is far less rigorous than that of Frankfurt (1971), it is continuous with his view that in acting autonomously our first-order desires are motivated by our second-order desires. If we identify with our second-order desires, we can be said to be “acting freely” à la Dworkin (1970).

²⁰ This is of course continuous with Berlin's (2002) ideas about positive and negative liberty.

²¹ I am drawing here on Davidson's (2001a, 2001b, 2001c, 2001d) philosophy of action, as well as Vargas's (2013, 246-249) notion of “moral ecology”, according to which agency is shaped by the “social terrain” in which agents act, that one's social identity and position can affect how and whether we can act.

²² This notion is continuous with Russell (1996, 24), and to some extent with Lukes (2005) where he claims, “A exercises power over B when A affects B in a manner contrary to B's interests” (27). However, where Lukes wants to avoid paternalism, I do not wish to dismiss it outright, as illustrated by GLUTTONOUS TODDLER.

over our own actions is different to how our actions affect others. One might construct a matrix, from low to high agency on the x-axis, and low to high power on the y-axis. Those with a high degree of agency likely have some power (not “low”), though one might imagine Thoreau at Walden Pond, or an ascetic completely withdrawn from society, controlling his own life very well but without affecting anyone else, and so falling rather low on the y-axis. Those scoring high on both axes (“high” power and “high” agency) might include our Supreme Court Justices from above, but perhaps others too, such as Jeff Bezos and Bill Gates. And within the context of the factory floor, the assembly line worker likely fits into the low power/low agency corner of our matrix, fitting the sparkplugs to the engine over and over again. But now we arrive at the important square of the matrix, those with *little agency and a high degree* of power. The unlikeliness of this scenario indicates that it is likely that agency is a necessary though insufficient condition for power. How could one who cannot author her own actions also control others? For instance, we might imagine Plato’s unwilling tyrant who has a great deal of power but no agency. Such an example of someone with a lot of power but no agency is fairly absurd, because agency is required to exercise power *at all*.

Where we have a great degree of power over how others are affected, we will also have a high degree of agency with regard to our own choices. We would not be able to cause effects that harm others if we ourselves were not in charge of our own decisions. Agency matters because it brings out the notion of a

free choice, a reasonable prerequisite of responsibility, where people are responsible for their exercise of power. So, just as we might say that a punishment should be proportional to the committed crime, we can similarly say that agency is proportional to degree of power. Who or what you are matters.²³

Have I relied on the dichotomy of individuals and groups in any substantive way? Where the literature's views on agency center on the metaphysical difficulties of the mereology of agency, they seem to run into strange artifacts. Reframing agency in terms of the power/agency continuum avoids these artifacts because it allows for gradients within a group, borne out by degrees of power, of *what we have control over*. Those who don't have *any* reasonable control over a situation cannot be said to have caused that situation, in which case we allow individuals within a group that are merely agentially epiphenomenal constituents. Their membership is causally inert to the outcome.²⁴ While we might suppose this has anti-structural consequences, it need not be so. It is simply that those who have very little power relative to other members of the group also have, relatively, less agency to cause the outcomes. Hence, one might reasonably suppose that an impoverished white person still has some unfair autonomy (in virtue of white privilege) not enjoyed by other

²³ There is broad application here from CEOs, boards of directors, nations, racial groups (whites and BIPOC), castes and classes, to assembly line workers and more.

²⁴ As with any gradient, there are near-limit cases. A person may have such a small degree of agency within a group that she has, functionally, no real agency at all. The factory worker could quit her job, but that may be the only economic option for her, and we could suppose feeding her family is also a reasonable responsibility. I'm agnostic about where that point is for everyone and anyone – but the contextualized nature of power and agency by affiliation to some group may provide an inroad to its location within the group.

aces, yet very little autonomy in other situations (e.g., she is a factory worker who installs the faulty sparkplug). Harm, agency, and power are contextual and gradient, which is simply a more accurate way of understanding causality than the individual/group notions of agency and their thin notions of harm.

2.3. Continuum III: Responsibility

So far, we have said that harm, power, and agency exist on continuums. We have said also that power requires agency. But having a greater or lesser degree of agency does not mean much on its own unless it is put to use. That we can do something does not mean that we will or must do something. But when we do put our agency to use, the continuums interactively come into play:

- (1) the more power we have, the more agency we have, and
- (2) the more agency we have, the greater our ability to propagate effects.

Now, let's specify "effects" as *harms*. The degree of harm we can propagate is proportional to the degree of agency we have. So, given (1) and (2):

- (3) the more power we have, the greater our ability to propagate harms.

That we can propagate harms justifies the move, then, to a discussion of responsibility, because we can be *held responsible for the harms we cause*. The lower we fall on a continuum of power, the less capacity we have to propagate harm, and the higher we fall on a continuum of power, the greater our capacity is to propagate harm. So, if we grant continuums I and II, we uncover that the degree

to which we are responsible corresponds to where we fall on a continuum of power.²⁵

Let's elaborate by returning to some abovementioned examples. In VEHICLE, for instance, the assembly line worker played a role in causing some harm, but it is disproportionate in relation to those who made the decisions to use the unreliable transmission, as the worker simply installed it. So, the worker's ability to harm is of a lower degree because she has a lower degree of power over what is done, whereas higher-ups' ability to harm is far greater, because they have a higher degree of power over what is done. In the case of SHOES, the shoe company exploitatively harms swaths of laborers. Though shoe-purchasers did not directly create the conditions in which the use of sweatshops was decided on, they play a role in perpetuating those conditions, and so also cause harm. So, consumers fall somewhere lower on the continuum as opposed to those who authored the decision to use sweatshops, given that they have the power to perpetuate the conditions where sweatshop labor is "justified" but did not themselves decide to set those conditions in motion. So, although consumers have a certain degree of power, they nevertheless harm to a lesser degree than those who made the ultimate decisions. Finally, and straightforwardly, in the

²⁵ Feinberg (1968) and Mellema (2006) have hinted at the notion of degrees of responsibility. The former by emphasizing "complicity" and "vicarious" action (675-685). The latter by emphasizing "qualifying actions". Neither focuses on power proper.

case of LAW, the Supreme Court Justices have an immense capacity to propagate harm, namely on the poor communities of San Antonio and beyond.²⁶

The degree to which we can propagate harms corresponds to the degree to which we are responsible for those harms. So:

(4) the greater our ability to propagate harms, the greater our responsibility.

Thus, we arrive at the continuum view of responsibility. Given (1), (2), and (4):

(5) the more power we have, the greater our responsibility.

Harm is about the quantity of negative consequences and their severity. How many are harmed and how harmed are they, and how does that harm correlate to our power? The paper mill officials who decide that toxic waste should be dumped into a river seem to warrant a far greater degree of responsibility given its impact on those who live by and depend on the river (not to mention the local ecosystem) than do the low-level employees who press the pulp and roll the sheets. The point is that the continuum of harm propagation meets a continuum of responsibility, both of which involve varying degrees of power.

Let's consider some implications of the more sophisticated notion of responsibility. Structurally, there are those within the advantaged group that bear more responsibility for structurally problematic outcomes. The Supreme Court of 1972 bears more responsibility for structural harm than perhaps any

²⁶ See, for instance, Sutton (2012). It may also be worth wondering that an impact is not just *agentially* captured – i.e., what we cause – but that it is also *temporally* captured – i.e., the effects of what we cause over time.

other discretely nameable group of individuals. It is not that we just eliminate the idea of group responsibility for structural outcomes, but that within certain groups some are clearly more responsible than others. Such a view has a practical consequence where we *recognize* that not all individuals may enjoy privilege in all aspects of life, in which case individuals with little autonomy in general are not disincentivized to reject out of hand the notion that they do enjoy some unfair advantages.

Yet there are other cases where, though one is a member of a group where *those with more power are responsible*, one is not “just responsible” for things that one quite literally has no functional control over. The U. S. did not evacuate all Afghans in time at the end of August 2021. Now, in what sense is Singapore responsible for that? Not at all. In what sense is a farmer in the cranberry bogs of Wisconsin responsible for that? *Probably* not at all (agnosticism may be our only option here). As power tapers off and approaches zero, there is a binary between responsible or not. There may not be a functional difference between Singapore and the cranberry farmer, meaning there was nothing that they could have done about it.²⁷

What of the explanatory power or parsimony of collective accounts? We do not have to go searching for the details for *each and every* actor within a group precisely because power exists in classes. That is, all those on the floor of the

²⁷ The reader may be surprised to find that I’ve invited a binary into my view, as I’ve made it a goal to do away with binaries in this dissertation. But this is an issue about which I’m willing to bite the bullet, because saying there is not a binary here would be absurd – that some who cannot make any functional difference *at all* are responsible.

assembly line who are not forepersons do not have a choice, they are a subset of the set of the larger corporation. The middle management constitute another set with a higher degree, culminating in a board of directors with the highest. We can do the same for nations. We can do the same for economic classes – certain sets of individuals enjoy the economic privilege (autonomy) to choose to purchase an electric car and some do not. There may be no excuse for Bill Gates (I do not know the status of his automobiles), some excuses whether totally compelling or not for the middle class, and every excuse in the world for the poor. Thus, the seeming messiness of a gradient view is easily resolvable by sets of classes of gradients, just as students get “A’s” and not 93.000001% or 93.000002%. The point is that we can get as fine-grained as we want, but we don’t have to.

2.4. Conscious Choice and Unawareness

So far, a lack of conscious choice and unawareness have not been explored in any detail. It may be objected that:

(1) Some do not consciously choose to cause harms in which case they cannot be held responsible.

(2) Some are unaware of the harms they cause in which case they cannot be held responsible.

Per (1), if we do not consciously choose to cause harm, then we are not responsible, because it is not something over which we had control. Case closed.

But a lack of conscious choice may not be relevant for some. That is, per (2), some

may not know the relevant facts of the situation. They haven't got a clue about the harms they caused. But shouldn't part of responsible agency be knowing what it is we choose as well as the relevant facts of our choice context?

First, consider the plausibility of ignorance for those with power as an excusing condition of responsibility. Let's say the higher-ups at the car factory were actually ignorant of the unreliability of the transmission and the CEO of the car wash just did not know about the defective carwax. Given their power, they really ought to have known *because* they have power, and we can hold them responsible *because* they should have known better. The more power we have, the more we need to be aware of our omissions. If we *commit* to not being aware, then we are willfully evading responsibility, and perhaps denying the very *idea* of responsibility. Equally seriously, we might also consider the fact that the Supreme Court ought to have known the structural outcomes they solidified were it the case that they pled ignorance. Finally, there are cases of *active ignorance*, where here all I mean is that someone whose responsibility it is to report wrongdoings actively goes out of her way to be unaware — one might imagine a police officer telling a friend that she can't hear talk of her friend's criminal behavior. There are all manner of cases where ignorance is not a valid excuse — and all of them involve a person in power *who ought to know better*.

Lastly, what of those whose power does not seem to play a role in their ignorance? Yet, consider Medina: a critical element of responsible agency is being “minimally knowledgeable about one's mind and one's life, about the social

world and the particular others with whom one interacts, and about the empirical realities one encounters” (2013, 127). Though his view falls within the purview of racism, it provides a lens through which we can see that we are none of us merely atomic individuals responsible for only ourselves, but that we are connected to others in such a way that our actions have a broad impact (133ff.). It seems not just that an individual must come to terms with this herself, but that she must also see that she is networked with others and with society writ large. The point is that thinking about what *I* have caused includes thinking about what *I* have affected. So, unless we were to deny the idea of moral responsibility, what we find is that in looking out for ourselves we are looking out, too, for other individuals and groups – those affected by the actions we cause.

Conclusion

Earlier in the chapter privilege was a springboard for talk about disproportionate power. After all, what is privilege but a kind of power, and what is power but a kind of privilege? My thinking is that if we take “privilege” in a broader sense than *white* privilege as is seen in the philosophy of race literature, then in the continuum style of thinking of this chapter, there are further depths to plumb where it comes to responsibility. That is, beyond the continuum framework’s providing nuance in terms of who should be *held to account* and by how much, it may provide nuance, also, to other senses of “responsibility”, for instance, in our being able to *take up* responsibility and our being able *to do* responsible actions. Outside privilege in the context of race, there

are degrees of privilege *in being able to be* responsible in other of life's arenas: as a consumer, in relation to the environment, and so on.

Consider the capacity to *take up* responsibility against the backdrop of privilege. If we should no longer buy clothing made in sweatshops or from manufacturers that leave a large carbon footprint, then it would seem that only those with a surplus of economic privilege are able to take responsibility for buying clothes from vendors that do not engage in those practices. Perhaps we buy an organic cotton farm and make our own shirts, thus avoiding sweatshop labor and leaving no carbon imprint. And those sufficiently motivated to mitigate harm to the environment who buy electric cars can do so given their economic capital. But some who are likewise motivated have little to no choice but to buy the beater car, thus perpetuating environmental harm. They simply do not have the means to buy the electric car.

This sort of thinking takes us to the capacity *to do* responsible actions. It should be eye-opening that those who have a lesser degree of privilege sometimes *cannot do* responsible actions – those without the means cannot buy the organic cotton farm or the electric car. So perhaps, then, those with great privilege who can buy the farm and the electric car owe more in terms of undertaking responsible actions. In either case, we come full circle from the capacity to *take up* responsibly and *to do* responsible actions to accountability. The privileged may well be accountable for their *lack* of shouldering the burden of responsibility. They bear a greater burden for society.

Conclusion

Charting Out New Paths: Two Sites for Future Research

Were this dissertation to be abridged, its major identifiable theme would be *philosophy as amelioration*. This theme ran through the foregoing pages both in terms of philosophical analysis and of attempting to improve our lives and practices. I've attempted, generally, to demonstrate amelioration's force on how both theory and *praxis* are conducted and construed in philosophical inquiry.

In chapter 2, I employed a re-tooled notion of "objectivity" that, at a minimum, means *general*. How did that improve theory in the philosophy of race and racism? The sense of "objective" I operated with coherently, non-reductively unified the ideas of the literature. This allowed for a general view of various perspectives in the literature, which in turn allowed for a basic idea of what we're talking about. And if the major views of racism can be unified, then we have a general method to ameliorate racism – theory in the service of *praxis*.

In chapter 3, I hope to have shown that we needn't be forced into theoretically accepting either a social-structural or individualist lens through which to understand racism. The upshot of doing so is that knowledgeable, anti-racist whites must always confront the reality that social forces play out in their individual minds – it is never safe to presume that one isn't individually racist nor that one is absolvable given the "transcendent" aspect of structural racism. I explored this idea in a psychologically realistic way via *overlooking*, where

ameliorating racism from the white perspective is an ongoing, but not hopeless, project – theory, again, in service of *praxis*.

Chapter 4, in a theoretical move similar to chapter 3, re-framed responsibility, arguing that we needn't be forced to choose between responsibility models divided into individual versus social camps. I argued that we should instead locate responsibility on a power continuum, which provided a realistic lens by which persons and groups are held to account without having to explicitly delimit ourselves to one or the other. In being more generally convincing, this might actually incentivize people to take responsibility where they might not otherwise – theory, once more, in service of *praxis*.

But beyond rehearsing this dissertation's path, the purpose of this conclusion is to indicate some new directions borne out by the dissertation's themes. I want to consider something new, and offer a future research agenda that isn't simply reducible to the ideas already presented. Where can those ideas go where they haven't already been? Permit me to sketch here two sites for further thinking, and why we should care about exploring them.

1. The Privilege of Responsibility

The first site for further thinking has to do with complicating the moral idea of responsibility. Chapter 4 presented a framework for responsibility which located it on a continuum of power. The idea was simple enough: the more power we have, the greater our responsibility. In that chapter I used privilege as a positive launchpad to explore power. But there is something more to explore

beyond power, something new if the conversation were re-directed *back* to privilege. What I'm interested in investigating is neither taking responsibility nor being held to account *for* one's privilege. Rather, the idea is to unpack the idea that responsibility *itself* is a privileged affair. I want to develop a general position according to which responsibility is asymmetrical, that it is a privilege disproportionately available to the public. My thinking is that such a position is supported by the material, socio-economic reality where privilege is involved in the capacity to *take up* responsibility *at all* for oneself, one's actions, and for others. Doing good – i.e., doing responsible actions – is generally available only to those who have the *means* to do good. A startling implication is that those with privilege are more praiseworthy than those without it.

1.1. The Privilege of Taking Up Responsibility

One project that falls out of this line of thinking is an examination of the privilege involved in being able to take up responsibility, that is, our *capacity* to be responsible. Those with more material resources are far better positioned to take responsibility, whereas those without resources may not have the capacity to take responsibility at all. The material capital we have in some way determines whether we can even be responsible, because it can limit our choices and the control over what we can do. Call this the *materiality of responsibility thesis*. To whom is the capacity open? Instead of answering such a question by turning to the metaphysics of agency, how a person must be composed to be a viable candidate for responsible agency, the idea is to turn away from metaphysics and

turn towards socio-economic conditions. We needn't look further than the current state of the world to see how this idea plays out. It is a luxury to work from home during the COVID-19 pandemic in order to mitigate the spread of the coronavirus. But of course, some do not have the luxury to take responsibility for themselves and their community in this case, having to continue work indoors amongst others. The point is that the ability to take responsibility is limited by privilege and, conversely, unbounded by it.

We should care about this because if it is true that there is real privilege in taking up responsibility, then there are several odd implications that challenge the very idea of responsibility and its practice. While it may be appealing to say that responsibility is symmetrical, that it is more or less the same across the board, it turns out that responsibility is *asymmetrical*. It is asymmetrical not only in the sense of chapter 4 where some have more responsibility than others, but also in the sense that some do not have the capacity, the material resources, to engage in responsibility practices. That is, some have very little capacity, if at all, to be responsible. And if one's capacity to be a responsible agent depends on one's privileged status, then our social practice of holding agents to account, blaming or praising them, is *also* privileged. We may not want to be so quick, for example, to say that *all* unvaccinated persons during the pandemic are blameworthy, as those in underserved predominately black and brown neighborhoods may not have the ability to get vaccinated. Perhaps an aspect of privilege is not asking the question, "*Which* unvaccinated persons are to blame?"

1.2. *The Privilege of Doing Good and of Praise*

Another project follows from the *materiality of responsibility thesis*. The thesis reveals that responsibility is a kind of commodity, that doing something responsible – doing good – is generally open only to those who can “afford” it. If the capacity to take up responsibility is privileged, then *doing* responsible actions is also a matter of privilege. Working from home during the pandemic, buying the electric car to mitigate environmental harm, shopping for clothes only at fair trade stores so as to not support sweatshop labor – all of these undertakings are available only to a sub-set of the population. So, *doing* some responsible actions does not admit of equal opportunity. Not all can work from home or buy an electric car or fair-trade clothing, and so on. The idea conjures up Socrates’ interrogation of Cephalus at the beginning of the *Republic*, the frustrated result of which is that it is fairly easy to be a good person, to do good things, if one is wealthy. But if doing responsible actions is a privilege, and we praise those who do responsible actions, then praise too is privileged. Call this the *privilege of praise thesis*, according to which if there is privilege involved in doing good, then the privileged are seemingly more deserving of praise, a peculiar and inequitable consequence.

We should find this consequence disquieting. If we think doing good deserves praise, but this is only a live option for the privileged among us, then the very idea of praise might rightly need re-framing as well. Should we think that those with a high degree of privilege are entitled to a high degree of praise?

If the privileged are sufficiently motivated enough to care about not supporting low-wage labor, they do not have to eat their meals at fast food restaurants or buy their books on Amazon.com, because they have the resources to source their food from the upscale organic food market and shop at the local bookstore. Our quotidian understanding of praise amounts to saying that Jeff Bezos and Bill Gates have perhaps the greatest capacity to be praiseworthy. It amounts to saying that the poor are less praiseworthy than the rich, that whites are more praiseworthy than other racialized groups, and so on. These are hard implications to swallow and should motivate a critical reappraisal of the material reality of our praising practices.

2. The White Philosopher and the Aggressor/Oppressor Perspective

The second site for further thinking has to do with the role of the white philosopher in the philosophy of race and racism and the aggressor/oppressor perspective. Chapters 2 and 3 provide the seeds for investigating these topics. In chapter 2, I raised the question of whether a white philosopher can be objective in the philosophy race, and whether she has a role in the area. There I hope to have established that it is possible to speak objectivity about racism in a way that respects first-person perspectives. In chapter 3, I pivoted to a discussion of racism from the perspective of the oppressor and showed what it is like to begin from a position of relative competence and yet watch racism nevertheless unfold. Can these seeds bear any further fruit?

Permit me to frame a general position on these matters by speaking autobiographically. What interests me, and what should be of interest to any white person in this field, is the perspective of the oppressor and whether a white philosopher has anything at all to say about race and racism, *because* that is my situation. Do I belong in this field? Is my perspective of any use? My thinking is that there is more to be learned about the moral psychology of white folk from the white perspective, a theoretical enterprise in the service of ameliorative results.

2.1. The White Philosopher Approaches Whiteness and White Racism

White philosophers have of course approached whiteness and white racism. That is not new. But what I want to draw attention to are constellations of views that are representatively expressed by the notion of *white fragility* (DiAngelo 2011, 2018), the discomfort and denial whites feel in the face of racial issues. While I do not think that white fragility as a phenomenon should be jettisoned, I do think that the approach that undergirds it is not one that is likely to change many hearts and minds of white folk. My suggestion is that views like white fragility are too caught up in theory-*not-in-the-service-of-praxis*.

Developing a view where white moral psychology is reduced to a state of constant vigilance about its own moral failings is likely to generate further white fragility. Perhaps that is the point. But the issue is that white fragility becomes something of a self-fulfilling prophecy, a “gotcha” theory of the white mind. Will this galvanize the (white) public? That seems unlikely. While whites are certainly

not morally exempt from putting in the work of self-awareness, there seems to be too little onus on the practical reality of something like white fragility, a simplistic psychology without much in the way of practical consequences.

The implication of thinking through white moral psychology in this way is that we should have very little faith in the amelioration of racism, that it may not be ameliorable at all. Even if we think the problem of racism will never be solved, we should not lose faith in amelioration as a kind of regulative ideal. In the spirit of chapter 2, getting descriptively clear *first* helps us prescribe more effectively. Approaching white psychology from a moralistic starting point, as I think white fragility does, likely does more to “turn off” whites than it does to make things better for the oppressed. Shouldn’t that be our goal? Thus, my thinking is that a more realistic psychology, with a non-moralistic starting point, might better serve the goal of amelioration. A neutral theoretical approach in the service of praxis, presenting white psychology convincingly, is a better strategy for persuading the public – it may get them “on board” for real change.

2.2. Microaggressions from the Perspective of the Aggressor/Oppressor

The question of what a white philosopher can reasonably do in the philosophy of race and racism may find an unexpected home in the literature on microaggressions, those verbal or nonverbal slights (intentional or not) which express negative messages to persons on the basis of their membership in a marginalized group. Against the backdrop of theory-in-the-service-of-praxis once again, how might the perspective of the white philosopher help here? That

is, how might the perspective of the oppressor, or in the case of microaggressions, the *aggressor* be of theoretical use for practically targeting microaggressions?

The white philosopher's role is of course limited. We cannot know completely what it is like to be the target of a microaggression from the standpoint of the oppressed. Neither can we speak to certain aspects of the harm that is incurred. We cannot honestly begin from the perspective of the victim, as we cannot prescribe the nature of others' experiences outside our race (this seems especially true in the case of white folk). So, what can the white philosopher do?

What I propose, in the spirit of overlooking, is to approach microaggressions from the perspective of the white aggressor, not because the victim's perspective is ignorable, but *because* I am white. There is nothing wrong, on its face, with taking up this perspective in the service of mitigating harms. And in fact, this seems the right sort of role for the white philosopher, because there is no pretending to have any perspective other than one's own. There is a legitimacy to an approach from the "inside" of whiteness that is not something to shy away from, but rather, something to be embraced – namely that what whites may consciously ascribe themselves to ("I'm an anti-racist"; "I'm 'woke' enough not to engage in microaggressions"; and so on) is different than what they really feel and perhaps talk about outside academic circles. To talk about overlooking is to offer an honest appraisal that whites with education, training, and a morally competent viewpoint really do think they aren't racist and that's the end of it.

They've put in all the relevant work. Now, they may not say this aloud, but that seems to be how it is from the "inside". So, overlooking matters because it calls that psychology to account.

It is from that neutral descriptive starting point, how white psychology really is, that we can begin to make things better. Of course, "neutral" is not meant deny that from the perspective of the oppressed, microaggressions are aggressive. This program is not for the sake of the oppressor. But oppression is a relation – if we want to relieve the oppressed, then the manifestations of how the oppressor thinks and oppresses is clearly relevant. The point is that if we don't look through the perspective of the aggressor, we'll miss out on a realism about her psychology, and what we get is a subtler, greater range of the phenomenon. In the style of overlooking, it is a diagnostic for those who don't want to engage in microaggressions from the perspective of the aggressor.

BIBLIOGRAPHY

- Ahmed, Sara. 2007. "A Phenomenology of Whiteness." *Feminist Theory* 8 (2): 149-168.
- Alcoff, Linda Martín. 1998. "What Should White People Do?" *Hypatia* 13 (3): 6-26.
- Alcoff, Linda Martín. 2006. *Visible Identities: Race, Gender and the Self*. New York: Oxford University Press.
- Alcoff, Linda Martín. 2007. "Epistemologies of Ignorance: Three Types." In *Race and Epistemologies of Ignorance*, edited by Shannon Sullivan and Nancy Tuana, 39-57. Albany, NY: State University of New York Press.
- Alcoff, Linda Martín. 2012. "Alien and Alienated." In *Reframing the Practice of Philosophy: Bodies of Color, Bodies of Knowledge*, edited by George Yancy, 23-43. Albany, NY: State University of New York Press.
- Allen, Theodore W. 2012a. *The Invention of the White Race*. Vol. 1, *Racial Oppression and Social Control*. London: Verso.
- Allen, Theodore W. 2012b. *The Invention of the White Race*. Vol. 2, *The Origin of Racial Oppression in Anglo-America*. London: Verso.
- Al-Saji, Alia. 2014. "A Phenomenology of Hesitation: Interrupting Racializing Habits of Seeing." In *Living Alterities: Phenomenology, Embodiment, and Race*, edited by Emily S. Lee, 133-172. Albany, NY: State University of New York Press.
- Amodio, David M. and Jillian K. Swencionis. 2018. "Proactive Control of Implicit Bias: A Theoretical Model and Implications for Behavior Change." *Journal of Personality and Social Psychology* 115 (2): 255-275.
- Arendt, Hannah. 2003. "Collective Responsibility." In *Responsibility and Judgement*, edited by Jerome Kohn, 147-158. New York: Schocken Books.
- Arendt, Hannah. 1972. "On Violence." In *Crises of the Republic*, 103-198. San Diego: Harcourt Brace & Company.
- Ayling, Pere. 2019. *Distinction, Exclusivity and Whiteness: Elite Nigerian Parents and the International Education Market*. Gateway East, Singapore: Springer.

- Bailey, Alison. 2007. "Strategic Ignorance." In *Race and Epistemologies of Ignorance*, edited by Shannon Sullivan and Nancy Tuana, 77-94. Albany, NY: State University of New York Press.
- Banaji, Mahzarin R. and Anthony G. Greenwald. 2016. *Blindspot: Hidden Biases of Good People*. New York: Bantam Books.
- Bell, Derrick. 1992. "Racial Realism." *Connecticut Law Review* 24 (2): 363-379.
- Berlin, Isaiah. 2002. *Liberty: Incorporating Four Essays on Liberty*. 2nd Edition. Edited by Henry Hardy. Oxford: Oxford University Press.
- Blair, Irene V., Jennifer E. Ma, and Alison P. Lenton. 2001. "Imagining Stereotypes away: The moderation of implicit stereotypes through mental imagery." *Journal of Personality and Social Psychology* 81 (5): 828-841.
- Blum, Lawrence. 2002a. *"I'm Not a Racist, But...": The Moral Quandary of Race*. Ithaca, NY: Cornell University Press.
- Blum, Lawrence. 2002b. "Racism: What It Is and What It Isn't." *Studies in Philosophy and Education* 21: 203-218.
- Bradley, Ben. 2012. "Doing Away with Harm." *Philosophy and Phenomenological Research* 2: 390-412.
- Bratman, Michael. 1999. *Face of Intention: Selected Essays on Intention and Agency*. New York: Cambridge University Press.
- Bratman, Michael. 2014. *Shared Agency: A Planning Theory of Acting Together*. Oxford: Oxford University Press.
- Brownstein, Michael and Jennifer Saul, eds. 2016. *Implicit Bias and Philosophy: Moral Responsibility, Structural Injustice, and Ethics*, vol. 2. Oxford: Oxford University Press
- Butcher, Jonathan and Mike Gonzalez. 2020. "Critical Race Theory, the New Intolerance, and Its Grip on America." *The Heritage Foundation*, December 7, 2020. <https://www.heritage.org/sites/default/files/2020-12/BG3567.pdf>.
- Charmichael, Stokely and Charles V. Hamilton. 1967. *Black Power: The Politics of Liberation in America*. New York: Random House.
- Corlett, J. Angelo. 2001. "Collective Moral Responsibility." *Journal of Social Philosophy* 32 (4): 573-584.

- Correll, Joshua, Charles M. Judd, Bernadette Park, Melody S. Sadler, and Bernd Wittenbrink. 2007. "Across the Thin Blue Line: Police Officers and Racial Bias in the Decision to Shoot." *Journal of Personality and Social Psychology* 92 (6): 1006-1023.
- Curry, Tommy J. 2007. "Please Don't Make Me Touch 'Em: Towards a Critical Race Fanonianism as a Possible Justification for Violence against Whiteness." *Democracy, Racism and Prisons* 5: 133-158.
- Darby, Derrick and Nyla Branscombe. 2014. "Beyond the Sins of the Fathers: Responsibility for Inequality." *Midwest Studies in Philosophy* XXXVIII: 121-137.
- Dasgupta, Nilanjana and Anthony G. Greenwald. 2001. "On the Malleability of Automatic Attitudes: Combating Automatic Prejudice With Images of Admired and Disliked Individuals." *Journal of Personality and Social Psychology* 81 (5): 800-814.
- Davidson, Donald. 2001a. "Actions, Reasons, and Causes." In *Essays on Actions and Events*, 3-20. Oxford: Oxford University Press.
- Davidson, Donald. 2001b. "Agency." In *Essays on Actions and Events*, 43-62. Oxford: Oxford University Press.
- Davidson, Donald. 2001c. "Intending." In *Essays on Actions and Events*, n-n. Oxford: Oxford University Press.
- Davidson, Donald. 2001d. "Freedom to Act." In *Essays on Actions and Events*, n-n. Oxford: Oxford University Press.
- Davis, Angela Y. 1983. *Women, Race & Class*. New York: Vintage Books.
- Davis, Angela Y. 1990. *Women, Culture, & Politics*. New York: Vintage Books.
- Devine, Patricia G., Forscher, A. J. Austin, and W. T. L. Cox. 2012. "Long-Term Reduction in Implicit Race Bias: A Prejudice Habit-Breaking Intervention." *Journal of Experimental Social Psychology* 48 (6): 1267-78.
- Dewey, John. 1922. *Democracy and Education*. New York: The MacMillan Company.
- Dewey, John. 1954. "An Added Note as to the 'Practical'." In *Essays in Experimental Logic*, 330-334. New York: Dover Publications, Inc.

- Dewey, John. 1958. *Experience and Nature*. New York: Dover Publications, Inc.
- Dewey, John. 2002. *Human Nature and Conduct*. Mineola, NY: Dover Publications, Inc.
- Dewey, John. 2004. *Reconstruction in Philosophy*. Mineola, NY: Dover Publications, Inc.
- DiAngelo, Robin. 2011. "White Fragility." *International Journal of Critical Pedagogy* 3 (3): 54-70.
- DiAngelo, Robin. 2018. *White Fragility: Why It's So Hard For White People To Talk About Racism*. Boston: Beacon Press.
- Du Bois, W. E. B. 1986. "Dusk of Dawn." In *W. E. B. Du Bois Writings*, edited by Nathan Huggins, 549-801. New York: The Library of America.
- Dworkin, Gerald. 1970. "Acting Freely." *Noûs* 4 (4): 367-383.
- Evans, Michele K., Lisa Rosenbaum, Debra Malina, Stephen Morrissey, and Eric J. Rubin. 2020. "Diagnosing and Treating Systemic Racism." *The New England Journal of Medicine* 383 (3): 274-276.
- Fanon, Frantz. 1952. *Black Skin, White Masks*. New York: Grove Press.
- Feagin, Joe R. 2006. *Systemic Racism: A Theory of Oppression*. New York: Routledge.
- Feinberg, Joel. 1968. "Collective Responsibility." *The Journal of Philosophy* 65 (21): 674-688.
- Feinberg, Joel. 1984. *Harm to Others*. Oxford: Oxford University Press.
- Feinberg, Joel. 1986. "Wrongful Life and the Counterfactual Element in Harming." *Social Philosophy and Policy* 4: 145-178.
- Fischer, John Martin and Mark Ravizza. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
- Foucault, Michel. 1978. *The History of Sexuality, Volume 1: An Introduction*. Translated by Robert Hurley. New York: Vintage Books.

- Foucault, Michel. 1980. *Power/Knowledge: Selected Interviews and Other Writings, 1972-77*. Edited by Colin Gordon. New York: Pantheon Books.
- Frankenberg, Ruth. 1993. *White Women, Race Matters: The Social Construction of Whiteness*. Minneapolis: University of Minnesota Press.
- Frankfurt, Harry. 1971. "Freedom of the Will and the Concept of a Person." *The Journal of Philosophy* 68 (1): 5-20.
- Frankish, Keith. 2010. "Dual-Process and Dual-System Theories of Reasoning." *Philosophy Compass* 5 (10): 914-926.
- French, Peter A. 1979. "The Corporation as a Moral Person." *American Philosophical Quarterly* 16: 207-217.
- French, Peter A. 1984. *Collective and Corporate Responsibility*. New York: Columbia University Press.
- French, Peter A. 1995. *Corporate Ethics*. Fort Worth, TX: Harcourt Brace.
- French, Peter A., editor. 1998. *Individual and Collective Responsibility*. Rochester, VT: Schenkman.
- Frye, Marilyn. 1992. *Willful Virgin: Essays in Feminism, 1976-1992*. Freedom, CA: The Crossing Press.
- Garcia, Jorge L. A. 1996. "The Heart of Racism." *Journal of Social Philosophy* 27 (1): 5-46.
- Garcia, Jorge L. A. 1999. "Philosophical analysis and the moral concept of racism." *Philosophy and Social Criticism* 25 (5): 1-32.
- Gawronski, Bertram, Roland Deutsch, Sawsan Mbirkou, Beate Seibt, and Fritz Strack. 2008. "When 'Just Say No' is not enough: Affirmation versus negation training and the reduction of automatic stereotype activation." *Journal of Experimental Social Psychology* 44 (2): 370-377.
- Glasgow, Joshua. 2009. "Racism as Disrespect." *Ethics* 120: 64-93.
- Goldberg, David Theo. 2021. "The War on Critical Race Theory." *Boston Review*, May 7, 2021. <http://bostonreview.net/race-politics/david-theo-goldberg-war-critical-race-theory>.

- Goldman, Alvin I. 1972. "Toward a Theory of Social Power." *Philosophical Studies* 23 (4): 221-268.
- Gordon, Lewis. 1995. *Bad Faith and Antiracist Racism*. Atlantic Highlands, NJ: Humanities.
- Hanser, Matthew. 2008. "The Metaphysics of Harm." *Philosophy and Phenomenological Research* 77 (2): 421-450.
- Hardin, Curtis D. and Mahzarin R. Banaji. 2013. "The Nature of Implicit Prejudice: Implications for Personal and Public Policy." In *The Behavioral Foundations of Public Policy*, edited by Eldar Shafir, 13-30. Princeton, NJ: Princeton University Press.
- Harris, Adam. 2021. "The GOP's 'Critical Race Theory' Obsession." *The Atlantic*, May 7, 2021.
<https://www.theatlantic.com/politics/archive/2021/05/gops-critical-race-theory-fixation-explained/618828/>.
- Harris, Leonard. 1999/2020. "What, Then, Is Racism?" In *A Philosophy of Struggle: The Leonard Harris Reader*, edited by Lee A. McBride III, 55-67. London: Bloomsbury Academic.
- Harris, Leonard. 2002/2020. "Insurrectionist Ethics: Advocacy, Moral Psychology, and Pragmatism." In *A Philosophy of Struggle: The Leonard Harris Reader*, edited by Lee A. McBride III, 175-187. London: Bloomsbury Academic.
- Haslanger, Sally. 2012. *Resisting Reality: Social Construction and Social Critique*. Oxford: Oxford University Press.
- Hoagland, Sarah Lucia. 2007. "Denying Relationality: Epistemology and Ethics and Ignorance." In *Race and Epistemologies of Ignorance*, edited by Shannon Sullivan and Nancy Tuana, 95-118. Albany, NY: State University of New York Press.
- Hobbes, Thomas. 1994. *Leviathan*. Indianapolis: Hackett Publishing Company, Inc.
- Holroyd, Jules. 2012. "Responsibility for Implicit Bias." *Journal of Social Philosophy* 43 (3): 274-306.
- hooks, bell. 1989. *Talking Back: Thinking Feminist, Thinking Black*. Boston: South End Press.

- Isaacs, Tracy. 2006. "Collective Moral Responsibility and Collective Intention." *Midwest Studies in Philosophy* XXX: 59-73.
- Isaacs, Tracy. 2014. "Collective Responsibility and Collective Obligation." *Midwest Studies in Philosophy* XXXVIII: 40-57.
- Jones, Richard D. 2009. "Philosophical Methodologies of Critical Race Theory." *Georgetown Journal of Law & Modern Critical Race Perspectives* 1 (1): 17-40.
- Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.
- Kawakami, Kerry, John F. Dovidio, Jasper Moll, Sander Hermsen, and Abby Russin. 2000. "Just say no (to stereotyping): Effects of training in the negation of stereotypic associations on stereotype activation." *Journal of Personality and Social Psychology* 78 (5): 871-888.
- Kelly, Daniel and Erica Roedder. 2008. "Racial Cognition and the Ethics of Implicit Bias." *Philosophical Compass* 3 (3): 522-540.
- Kim, David Haekwon. 2020. "Racist Love." In *50 Concepts for a Critical Phenomenology*, edited by Gail Weiss, Ann V. Murphy, and Gayle Salamon, 295-302. Evanston, IL: Northwestern University Press.
- Klocksien, Justin. 2012. "A Defense of the Counterfactual Comparative Account of Harm." *American Philosophical Quarterly* 48: 285-300.
- Koopman, Colin. 2006. "Pragmatism as a Philosophy of Hope: Emerson, James, Dewey, Rorty." *Journal of Speculative Philosophy* 20 (2): 106-116.
- Krasne, Ellie. 2020. "How Leftists' Critical Race Theory Poisons Our Discussion of Racism." *The Heritage Foundation*, June 29, 2020.
<https://www.heritage.org/civil-society/commentary/how-leftists-critical-race-theory-poisons-our-discussion-racism>.
- Lee, Emily S, editor. 2014. *Living Alterities: Phenomenology, Embodiment, and Race*. Albany, NY: SUNY Press.
- List, Christian and Philip Pettit. 2011. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press.
- López, Ian Haney. 2006. *White by Law: The Legal Construction of Race*. New York: New York University Press.

- Lukes, Steven. 2005. *Power: A Radical View*. 2nd Edition. London: Palgrave Macmillan.
- Mackenzie, Catriona. 2018. "Moral Responsibility and the Social Dynamics of Power and Oppression." In *Social Dimensions of Moral Responsibility*, edited by Katrina Hutchison, Catriona Mackenzie, and Marina Oshana, 59-80. Oxford: Oxford University Press.
- MacMullan, Terrance. 2009. *Habits of Whiteness: A Pragmatist Reconstruction*. Bloomington, IN: Indiana University Press.
- MacMullan, Terrance. 2015. "Facing Up to Ignorance and Privilege: Philosophy of Whiteness as Public Intellectualism." *Philosophy Compass* 10 (9): 646-60.
- Madva, Alex. 2016. "A Plea for Anti-Anti-Individualism: How Oversimple Psychology Misleads Social Policy." *Ergo* 3 (27): 701-728.
- Mathiesen, Kay. 2006. "We're All in This Together: Responsibility of Collective Agents and Their Members." *Midwest Studies in Philosophy* XXX: 240-255.
- May, Larry. 1992. *Sharing Responsibility*. Chicago: University of Chicago Press.
- McBride III, Lee A. 2013. "Insurrectionist Ethics and Thoreau." *Transactions of the Charles S. Peirce Society* 49 (1): 29-45.
- McBride III, Lee A. 2017. "Insurrectionist Ethics and Racism." In *The Oxford Handbook of Philosophy and Race*, edited by Naomi Zack, 225-234. Oxford: Oxford University Press.
- McKenna, Michael. 2018. "Power, Social Inequities, and the Conversational Theory of Moral Responsibility." In *Social Dimensions of Moral Responsibility*, edited by Katrina Hutchison, Catriona Mackenzie, and Marina Oshana, 38-58. Oxford: Oxford University Press.
- Medina, José. 2013. *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imaginations*. Oxford: Oxford University Press.
- Mills, Charles W. 1997. *The Racial Contract*. Ithaca, NY: Cornell University Press.
- Mills, Charles W. 2003. "'Heart Attack': A Critique of Jorge Garcia's Volitional Conception of Racism." *The Journal of Ethics* 7 (1): 29-62.

- Mills, Charles W. 2007. "White Ignorance." In *Race and Epistemologies of Ignorance*, edited by Shannon Sullivan and Nancy Tuana, 11-38. Albany, NY: State University of New York Press.
- Mills, Charles W. 2012. "Philosophy Raced, Philosophy Erased." In *Reframing the Practice of Philosophy: Bodies of Color, Bodies of Knowledge*, edited by George Yancy, 45-70. Albany, NY: State University of New York Press.
- Mellema, Gregory. 2006. "Collective Responsibility and Qualifying Actions." *Midwest Studies in Philosophy* XXX: 168-175.
- Nussbaum, Martha C. 1988. "Non-Relative Virtues: An Aristotelian Approach." *Midwest Studies in Philosophy* 13 (1): 32-53.
- Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. 2019. "Dissecting racial bias in an algorithm used to manage the health of populations." *Science* 366: 447-453.
- Omi, Michael and Howard Winant. 1994. *Racial Formation in the United States From the 1960s to the 1990s*. Second Edition. New York: Routledge.
- Ortega, Mariana. 2006. "Being Lovingly, Knowingly Ignorant: White Feminism and Women of Color." *Hypatia* 21 (2): 56-74.
- Oshana, Marina. 2018. "Ascriptions of Responsibility Given Commonplace Relations of Power." In *Social Dimensions of Moral Responsibility*, edited by Katrina Hutchison, Catriona Mackenzie, and Marina Oshana, 81-109. Oxford: Oxford University Press.
- Outlaw (Jr.), Lucius T. 2007. "Social Ordering and the Systematic Production of Ignorance." In *Race and Epistemologies of Ignorance*, edited by Shannon Sullivan and Nancy Tuana, 197-211. Albany, NY: State University of New York Press.
- Peruche, B. Michelle and E. Ashby Plant. 2006. "The Correlates of Law Enforcement Officers' Automatic and Controlled Race-Based Responses to Criminal Suspects." *Basic and Applied Social Psychology* 28 (2): 193-199.
- Radzik, Linda. 2001. "Collective Responsibility and Duties to Respond." *Social Theory and Practice* 27 (3): 455-471.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.

- Rich, Adrienne. 1979. *On lies, secrets, and silence: Selected prose 1966–1978*. New York: W. W. Norton & Company.
- Roediger, David R. 1991. *The Wages of Whiteness: Race and the Making of the American Working Class*. London: Verso.
- Roediger, David R. 2002. *Colored White: Transcending the Racial Past*. Berkeley: University of California Press.
- Russell, Bertrand. 1996. *Power: A New Social Analysis*. London: Routledge.
- Russell, Camisha A. 2018. *The Assisted Reproduction of Race*. Bloomington, IN: Indiana University Press.
- Sadler, Brook Jenkins. 2006. "Shared Intentions and Shared Responsibility." *Midwest Studies in Philosophy* XXX: 115-144.
- Sawchuk, Stephen. 2021. "What is Critical Race Theory, and Why Is It Under Attack?" *Education Week*, May 18, 2021.
<https://www.edweek.org/leadership/what-is-critical-race-theory-and-why-is-it-under-attack/2021/05>.
- Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Schmid, W. Thomas. 1996. "The Definition of Race." *Journal of Applied Philosophy* 13 (1): 31-40.
- Schwartz, Robert. 2015. *Rethinking Pragmatism: From William James to Contemporary Philosophy*. Malden, MA: Wiley Blackwell.
- Shelby, Tommie. 2002. "Is Racism in the 'Heart'?" *Journal of Social Philosophy* 33 (3): 411-420.
- Silva, Grant. 2019. "Racism as Self-Love." *Radical Philosophy Review* 22 (1): 85-112.
- Silver, David. 2002. "Collective Responsibility and the Ownership of Actions." *Public Affairs Quarterly* 16 (3): 287-304.
- Smart, J. J. C. 1961. "Free-Will, Praise and Blame." *Mind* 70 (279): 291-306.
- Smart, J. J. C. 1973. "An Outline of a Utilitarian System of Ethics." In *Utilitarianism: For and Against*, edited by J. J. C. Smart and Bernard Williams, 3-74. Cambridge: Cambridge University Press.

- Smiley, Marion. 2017. "Collective Responsibility." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta.
<<https://plato.stanford.edu/archives/sum2017/entries/collective-responsibility/>>.
- Smith, Angela. 2005. "Responsibility for Attitudes: Activity and Passivity in Mental Life." *Ethics* 115 (2): 236-271.
- Sripada, Chandra. 2016. "Self-Expression: A Deep Self Theory of Moral Responsibility." *Philosophical Studies* 173 (5): 1203-1232.
- Stanovich, Keith E. and Richard F. West. 2000. "Individual differences in reasoning: Implications for the rationality debate?" *Behavioral and Brain Sciences* 23: 645-726.
- Stickers, Kenneth W. 2014. "...But I'm Not Racist": Toward a Pragmatic Conception of Racism." *The Pluralist* 9 (3): 1-17.
- Strawson, P. F. 1974. *Freedom and Resentment and Other Essays*. London: Routledge.
- Sullivan, Shannon. 2006. *Revealing Whiteness: The Unconscious Habits of White Privilege*. Bloomington, IN: Indiana University Press.
- Sutton, Jeffrey S. 2008. "San Antonio Independent School District v. Rodriguez and its Aftermath." *Virginia Law Review* 94: 1963-1986.
- Talbert, Matthew. 2019. "Moral Responsibility." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N Zalta.
<<https://plato.stanford.edu/archives/win2019/entries/moral-responsibility/>>.
- Taylor, Paul C. 2007. "Race, Ethics, Seduction, Politics: On Shannon Sullivan's Revealing Whiteness." *Journal of Speculative Philosophy* 21 (3): 201-209.
- Taylor, Paul C. 2016. *Black is Beautiful: A Philosophy of Black Aesthetics*. Malden, MA: John Wiley & Sons.
- Thaler, Richard H. and Cass R. Sunstein. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New York: Penguin Books.
- Thomson, Judith Jarvis. 2011. "More on the Metaphysics of Harm." *Philosophy and Phenomenological Research* 82: 436-458.

- Urquidez, Alberto G. 2020. *(Re-)Defining Racism: A Philosophical Analysis*. Cham, Switzerland: Palgrave MacMillan.
- Vargas, Manuel. 2013. *Building Better Beings: A Theory of Moral Responsibility*. Oxford: Oxford University Press.
- Vice, Samantha. 2010. "How Do I Live in This Strange Place?" *Journal of Social Philosophy* 41 (3): 323-342.
- Watson, Gary. 1996. "Two Faces of Responsibility." *Philosophical Topics* 24 (2): 227-248.
- Weber, Max. 2013. *Economy and Society*, vol. 1. Edited by Guenther Roth and Claus Wittich. Berkeley: University of California Press.
- Williams, Bernard. 1993. *Shame and Necessity*. Berkeley, CA: University of California Press.
- Wittgenstein, Ludwig. 1958. *The Blue and Brown Books*. New York: Harper & Row, Publishers.
- Wolf, Susan. 1987. "Sanity and the Metaphysics of Responsibility." In *Responsibility, Character, and the Emotions*, edited by Ferdinand Schoeman, 46-62. Cambridge: Cambridge University Press.
- Wolf, Susan. 1990. *Freedom Within Reason*. New York: Oxford University Press.
- Yancy, George. 2008. *Black Bodies, White Gazes: The Continuing Significance of Race*. New York: Rowman & Littlefield Publishers, Inc.
- Yancy, George. 2012. *Look, a White! Philosophical Essays on Whiteness*. Philadelphia: Temple University Press.
- Young, Iris Marion. 2004. "Responsibility and Global Labor Justice." *The Journal of Political Philosophy* 12 (4): 365-388.
- Young, Iris Marion. 2006. "Responsibility and Global Justice: A Social Connection Model." *Social Philosophy and Policy* 23: 102-130.
- Young, Iris Marion. 2011. *Responsibility for Justice*. Oxford: Oxford University Press.

Zimmerman, Michael J. 1985. "Sharing Responsibility." *American Philosophical Quarterly* 22 (2): 115-122.

Zimmerman, Michael J. 2015. "Varieties of Moral Responsibility." In *The Nature of Moral Responsibility: New Essays*, edited by Randolph Clarke, Michael McKenna, and Angela M. Smith, 45-64. Oxford: Oxford University Press.