THAT'S (ALSO) RACIST! ENTITY TYPE PLURALISM, RESPONSIBILITY, AND LIBERATORY NORMS

by

Lacey J. Davidson

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Philosophy

West Lafayette, Indiana

August 2019

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Daniel Kelly, Chair

　　Department of Philosophy

Dr. Leonard Harris

　　Department of Philosophy

Dr. Taylor Davis

　　Department of Philosophy

Dr. Ron Mallon

　　Department of Philosophy, Washington University in Saint Louis

Approved by:

　　Dr. Christopher Yeomans

　　　　Head of the Graduate Program

With all those engaged in the constant struggle for liberation

ACKNOWLEDGMENTS

Thank you to Indiana Voices of Women, Conversation Circles, Civic Engagement and Leadership Development at Purdue, and especially the Younger Women's Task Force for helping me find my place in Indiana and getting me connected to the movements I care most about. In addition, the epistemic norm interventions discussed in Chapter 5 come from my participation in these organizations. Thank you for being comrades and friends. Thank you especially to Melissa Gruver, Vanessa Pacheco, Jesus Romero, Barbara Kerkhoff, Christopher Munt, Alex Rusco, Emily Rames, Lyndsy Manz, Ulyssa Hester, Samira Fatemi, Kirsten Gibson, and Jessica Joneson.

Thank you to my friends not already mentioned, Britt Currie, Leah Fortgang, Hope Gulker, Quincy Law, Gabriel Lonsberry, for supporting me. A special thanks to Johanna Wu for assisting me with administrative and coordination issues. And a huge thanks to my writing partner, Will Penman, for making the writing process far more enjoyable. Thank you to my friends at Yoga Balance, Shari Gass, Lisa Nail, and Debra Steinhauer; its been great to teach yoga together.

Thank you to my childhood friends and friends from Springfield, OH. Thank you to Catie Stipe, Tori Amicon, Lauren Bender, Mary Dowling, Laura Nagel, and Tessa Donnell. Thank you especially to Rahsaad Jones who put up with me even when I didn't know anything but thought I did.

Finally, thank you to my family for raising me to pursue my interests and to work hard. Thank you, Mom, for teaching me to work for what is right even when it is not easy, and thank you, Dad, for teaching me about working hard, consistency, and contentment. Thank you, Sandy and Jay, for being role models to me and blazing the path.

As I often do, I'm sure I've forgotten someone in these acknowledgements. Please know that this is a reflection on me and not you. Thank you for giving me the support I needed to write this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Author: Davidson, Lacey, J. PhD
Institution: Purdue University
Degree Received: August 2019
Title: That's (Also) Racist! Entity Type Pluralism, Responsibility, and Liberatory Norms
Committee Chair: Daniel Kelly

Some philosophers (Blum 2002 and Anderson 2010) have argued that 'racism' and 'racist' have been used so widely that they have lost their conceptual potency and are no longer effective moral evaluations. For this reason, they think we should use other terms to identify racial injustices. It is the goal of this dissertation to argue against this conclusion. In Chapter 2, I develop tools for diagnosing the individualist versus structuralist debate within philosophical accounts of racism. I use these tools to show that both individualists and structuralists are committed to entity type monism or the view that only certain kinds of entities can be racist. I reject this view and argue for entity type pluralism. In Chapter 3, I move from entity type pluralism to develop an account of the application conditions for the predicate 'racist' that tell us when and why we should apply the predicate to particular entities. These two chapters serve to clarify RACISM. In Chapter 4, I develop new resources for understanding moral responsibility for racism, specifically for how agents can be held accountable for intervening upon racist non-agential entities like norms, policies, and institutions. I call these resources "oblique blame" and "intervention-sensitive moral responsibility." Intervention-sensitive moral responsibility gives way to a problem. Given the ways in which our current epistemic practices exclude the testimony of People of Color, we will have a hard time knowing when we are responsible in this intervention-sensitive way. I call this the Knowledge Problem. In Chapter 5, I bring together the literature on epistemic oppression and the empirically-informed norms literature to show that interventions into epistemic norms help solve this problem. I provide four candidate norms from activist and organizing communities as examples. Taken together, this dissertation shows that we need not discontinue our use of 'racism' and 'racist' and that the terms can be used effectively to hold each other accountable toward anti-racist aims and a liberated future.
.

CHAPTER 1.   INTRODUCTION

But the fact is, the word "racism" is too simplistic, too general, and too easy. You can use the word and not say much, unless the term is explained or clarified. – doris davenport (1981)

1.1   That's Racist!

According to KnowYourMeme.com, "'That's racist!' is an expression often used in jest to point out the politically incorrect or racially insensitive nature of a post or comment online."[1] The most popular form of this meme is a clip of a child saying, "That's racist!" on the sketch comedy show *Wonder Showzen*, but there are now many more versions in existence. Though the meme is sometimes used jokingly by those hoping to point out "over sensitivity" or to critique the pressure to be "politically correct," the primary use of the meme is to call out comments or behaviors that have questionable moral standing because of their use of racial stereotypes or essentializing beliefs about racial groups. In other words, the meme is used to hold people accountable for what they say online in a humorous, somewhat lighthearted way.[2]

Figure 1 "That's Racist!" Meme

---

[1] https://knowyourmeme.com/memes/thats-racist

[2] Some people have argued that the use of GIFs and memes of Black people by white people to express their emotions is a kind of "digital Blackface" (see https://www.teenvogue.com/story/digital-Blackface-reaction-gifs for commentary). My purpose here is not to suggest that everyone should use the "That's Racist!" meme (I prefer just the text without the GIF, myself), but that people (with many racialized identities) *do* use the meme as a humorous moral evaluation to various degrees of effectiveness.

However, there is some controversy in and outside of academia around whether calling someone or something racist takes things too far or is incorrect. In online communities, the primary worries seem to be censorship and the loss of a sense of humor (e.g. "we can't even make jokes any more without the sensitive snowflakes getting their feelings hurt!"). Within philosophy, the primary worry is around conceptual conflation, incoherency, and the loss of evaluative force (see especially, Blum 2002 and Anderson 2010). Within these critiques, there's a suggestion that we turn away from using the concept of RACISM and the predicate 'racist' in our discussions of social problems. It is my purpose in this dissertation to push back against these worries. It is my view that these terms and the ideas they express are (or can be) fundamental to anti-racist understandings that undergird action, and that, as doris davenport suggests, further exploration, explanation, and the deepening of clarity will be useful. As such, my goal in this dissertation is to build clarity around the concept of RACISM, especially as it relates to points of agreement and disagreement between accounts and the predicate 'racist'. I do this so the moral force of the predicate 'racist' can be preserved in a way that allows us to hold each other accountable.

In this Introduction, I'll talk a little bit about why I think the preservation of these terms, along with their moral force, is integral to the anti-racist project, and then expand on contemporary critiques of the terms. Then, I'll give an overview of the four main chapters of the dissertation and mention formatting conventions. Finally, I'll end with a few limitations of the project as it is currently constructed.

1.2 Why 'Racist' is Important

In 2010, I was sophomore at Wittenberg University in Springfield, OH. My "Knowing Bodies" seminar had just finished reading Yancy's (2008) *Black Bodies, White Gazes* when he visited our campus. In a presentation to our class, Yancy introduced us, and to me for the first time, to the concept of *racist anti-racists*. He argued that white people were always racist in virtue of their subject position in this historically contingent social reality, but that white people could also be anti-racists, folks committed to robust social change and the mitigation and elimination of racism. At this time, I considered myself someone committed to racial justice, but I didn't really think of myself as either a racist or an anti-racist, and it was after Yancy's presentation that I began to develop my identity as both. Thinking of myself as racist anti-racist has been fundamental to my understanding of myself as a philosopher and as a community organizer and activist.

I'm not alone in this experience. In her analysis of thirty women and the development of their anti-racist white feminist [3] identities, Frankenberg (1993, 162) emphasizes that the acknowledgement of one's own attitudes or actions as racist was important for the development of identity. For example, one woman says (1993, 162), "More [of my] friends than not tend to believe that they grew up with no racist attitudes whatsoever. Which may or may not be true. I'd like to believe that about myself, but I can't, because it's not true." For this white woman, understanding her own attitudes and actions as racist was integral to her development as an anti-racist feminist. Further, Frankenberg (1993) notes that as anti-racist identities developed, the women in her study were less likely to need to separate the "bad" racists from the "good" non-racists (170), and rather began to see themselves as both racist and anti-racist. This is an important conceptual shift because it shifts the focus from being "non-racist" or "good" to being an avowed anti-racist, which is a far stronger commitment, a commitment more likely to lead to genuine solidarity.

White women who are anti-racist also experience guilt and shame around their white identities, as well as fears of "appearing racist" in interactions with People of Color (Linder 2015). Using narrative exploration (journaling) with six white undergraduate women who identified as white anti-racist feminists, Linder (2015) explores the ways in which authentic relationships across racial lines are essential to the development of an anti-racist white feminist identity, but that experiences of guilt, shame, and fear often serve as a barrier to authentic relationships with People of Color and the social and political engagement that is essential to the anti-racist project. We see that the women in Linder's study tend to focus on their own experiences and are still fearful of being "bad" racists (also present in Frankenberg 1993, 170). In contrast to the women in Frankenberg's study, the six women in Linder's study do not talk about their own behaviors and actions as racist. Though more empirical work would need to be done to show this connection explicitly, it seems highly plausible that the hyper-focus on being a "good white person" may stem from the inability to see one's self as already racist. In other words, the development of one's own identity as racist *alongside* one's identity as anti-racist may play an important role in effectively engaging in the anti-racist project.

---

[3] 'White feminist' here refers to feminists who are white (who often recognized the problems within the feminist movement), not the contemporary use of 'white feminist' that is used to point out problematic forms of feminism that do not include a racial analysis in their accounts of gender-based oppression.

Perhaps most obviously, the term 'racism' and the predicate 'racist' are important to our notions of anti-racism or being an anti-racist. Without the positive notion of racism or what it is to be racist, we are unable to conceptualize the negation. This is highlighted through an activity[4] used by organizations with which I'm involved in anti-racist education. The facilitator draws four boxes on the board labeled "active racism," "passive racism," "active anti-racism," and "passive non-racism" (with the last category meaning to represent the absence of racism rather than an active commitment toward anti-racism). The group then suggests behaviors for each box (e.g. overtly racist comments, subtle reactions to People of Color, etc.). As the activity progresses, all of the boxes fill up with the exception of "passive non-racism." For every suggestion for this box, someone comes up with a reason why this action is either "active racism" or "passive racism," rather than non-racism. This activity is meant to highlight that living in a world such as ours doesn't allow complacency with respect to racism and that active anti-racism or taking up the identity and actions of an avowed anti-racist is the available option. If this demonstration reveals something true about the nature of our world, and I think it does, then we need the concepts of 'racism' and 'racist' to define ourselves against. We need to have a concept of what we won't tolerate in order to decide what is worth fighting for.

1.3   Getting Rid of 'Racist'

My purpose in this introduction thus far has been to show that the predicate 'racist' is being used effectively as a moral evaluation both when directed at others and when directed at oneself, especially as it relates to the development of anti-racist identities and organizational strategies. However, there has been some worry both in and outside of philosophy about conceptual inflation (Miles 1993, Blum 2002).[5] For example, the meme I showcased above is used in some cases to *avoid* important conversations around race. In other words, someone may direct it at themselves after saying something that may be perceived as racist in order to avoid an actual conversation or real accountability for what they've said or done. Further, the phrase and meme are sometimes

---

[4] I'm uncertain of the origins of this activity, but Mel Gruver learned it from Beverly Tatum and one version is available from the Tri-County Domestic & Sexual Violence Intervention Network in their Anti-Oppression Training and can be found here: https://www.pcc.edu/illumination/wp-content/uploads/sites/54/2018/05/active-passive-racism-anti-racism-exercise.pdf.

[5] For an example from popular culture, see this (underdeveloped) NPR clip: https://www.npr.org/2011/06/27/137451481/thats-racist-how-a-serious-accusation-became-a-commonplace-quip

used in cases in which there is clearly not anything racist going on (e.g. sorting laundry into dark and light loads as seen on *Parks and Recreation*) presumably to highlight oversensitivity.

Similarly, Blum (2002) worries that we've begun to overuse the terms 'racism' and 'racist' and that the result is that the terms have lost their moral potency. He writes (2002, 205):

> Some feel that the word is thrown around so much that anything involving 'race' that someone does not like is liable to castigation as 'racist'—for example, merely mentioning someone's race (or racial designation), using the word 'Oriental' for Asians without recognizing its origins and its capacity for insult, or socializing only with members of one's own racial group.

From this observation, he claims (2002, 203) that the "overuse" of terms like 'racism' and 'racist' now "constitute obstacles to understanding interracial dialogue about racial matters." Blum locates the primary problem as conceptual inflation. He thinks that we have come to use the terms 'racism' and 'racist' in such a broad way that they lose conceptual coherency (it is difficult to parse out exactly what is and is not racism) and their moral-evaluative force (because of the conceptual incoherency, calling someone a racist doesn't do anything, it doesn't change their behaviors or the behaviors of others). Due to these issues, Blum argues that we ought to limit the use of the terms 'racism' and 'racist'.

Though Blum notes that conceptual and definitional work has challenges, he also thinks that we can get to the "core meaning" of racism by looking at its historical use. In other words, we can find a limited concept of racism that doesn't fall prey to conceptual inflation. Blum (2002, 210) reports that the term 'racism' was first used by German sociologists to explain the ideology of racial superiority developed through Nazism and was then extended to explain things like segregation, South African apartheid, and European colonialism. From these uses, Blum argues that we can identify two core themes: inferiorization and antipathy. He then argues that racism essentially involves these representations or attitudes, which leads him (2002, 206) to conclude that "not all racial incidents are racist incidents" (206). Due to these commitments, Blum is considered to be an individualist or violitionalist in that he focuses on what is in the heads of people (their beliefs or attitudes toward other racial groups). This will be discussed further in Chapter 2, but for now the main lesson is that he seeks to limit the use of the term.

Like Blum, Anderson (2010, 5) has claimed that the concept of racism is "incoherently lumpy," and in her book *The Imperative of Integration* seeks to put forth an institutional-level change that might mitigate racial inequality. Importantly, her arguments do not rely on a robust

concept of racism. In fact, she (2010, 48) suggests: "…let us reserve 'racism' for judgments of serious vice, while observing that not all injustice is caused by a vicious character." So, though Anderson is typically characterized as a structuralist with respect to social explanation around race, she too is an individualist with respect to racism and further seeks to limit the use of the term.

In her critique of white feminism,[6] doris davenport ([1981]2015) argues that dominant (white) feminist spaces are diseased by racism and that deep (not surface) work needs to be done to "cure" these spaces. However, she too seems to think that the term 'racism' has some conceptual issues. She writes, "But the fact is, the word 'racism' is too simplistic, too general, and too easy. You can use the word and not say much, unless the term is explained or clarified. Once that happens, racism looks more like a psychological problem (or pathological aberration) than an issue of skin color" (81-82). She goes on to argue that the racism of white feminists is a psychological protection mechanism that is a result of white women's own oppression by men. Though davenport's discussion in this piece is intentionally limited to a critique of white feminist spaces, again we see the limiting of racism to describe an ideology or psychological problem. Though I will ultimately argue against this limiting, I will bring forth an essential lesson from davenport's work that the concept of racism must be clarified in order to preserve its normative force.

I'm not alone in thinking that we can effectively use the concept RACISM and the predicate 'racist' to move forward and that the limiting of our use is a dangerous strategy that can keep us from identifying social ills and dampen our anti-racist commitments. In discussing what he calls the "racist card," or the "But I'm not a racist" card that white people pull, Coates (2008) writes:

> In some measure, the narrowing of racism is an unfortunate relic of the civil rights movement, when activists got mileage out of dehumanizing racists and portraying them as ultra-violent Southern troglodytes. Whites may have been horrified by the fire hoses and police dogs turned on children, but they could rest easy knowing that neither they nor anyone they'd ever met would do such a thing. But most racism—indeed, the worst racism—is quaint and banal.

In direct contrast to Blum, Coates argues that historical use reflects a strategic move that does not necessarily track what is and is not racist and the idea that racist people or actions are exceptions, rather than the rule, is something that white people hide behind to avoid accountability.

---

[6] Here the term is used to identify the pernicious form of white feminism—not just feminists who are white, but an exclusionary white feminism that profits from the maintenance of white spaces.

Not only do we use the predicate 'racist' to describe things in the world and to morally evaluate each other (as with the meme with which I started), but there also seems to be good reasons to preserve this use rather than leave it behind. There are worries both inside and outside of the academy that no one knows what we are talking about anymore. However, this problem can be solved by increasing the conceptual clarity of the concept of RACISM, especially the predicate 'racist'. This increased clarity will allow us to use the predicate to hold each other accountable and build a different world with less suffering and tragic death of People of Color.

1.4   Overview of Chapters

In Chapter 2, develop a new framework called *levels of pluralism* for diagnosing points of agreement and disagreement between accounts of racism. Within this frame there are four sets of nested questions that accounts of racism address. I argue that these questions are in a dependence relationship with one another, such that taking a pluralist or monist position with respect to one level can influence one's position on other levels. I develop this frame to shed light on the debate between individualists and structuralists. Specifically, I argue that individualists and structuralists both (implicitly) ascribe to *entity type monism*, or the view that only certain kinds of entities (individuals or structures) can be the fundamental bearers of the predicate 'racist'. I argue that this shared commitment has limited our discussion of racism in ways that are both unnecessary and damaging to anti-racist aims. My primary aim in this chapter is to provide a pathway for going beyond the individualist versus structuralist debate.

In Chapter 3, I move forward from this entity type pluralism to develop an account of the application conditions for the predicate 'racist'. Though it is often not explicit, accounts of racism do include these applications, but making them explicit, taking time to show why and when we should apply the predicate to entities, increases the clarity around RACISM and the predicate 'racist' in ways that is helpful for the anti-racist project and the worries about conceptual inflation discussed above. I move forward from Harris' (2018) descriptive strategy and account of necro-being to articulate a conception of the predicate 'racist' as appropriately applied to entities that are involved in the violation of the right not to be killed. The application conditions, alongside the entity type pluralism that I developed in Chapter 1, show that the predicate 'racist' can be applied to many types of entities, including non-agential entities such as mental states, objects, social imaginaries, and institutional features like practices, policies, and laws.

In Chapter 4, I discuss the ways in which the predicate 'racist' is used as a description and as a moral evaluation. When we deploy 'racist' to describe an entity, we get a moral evaluation for free. However, the resources developed within the philosophical moral responsibility literature do not adequately capture the nature of these moral evaluations. To address this gap, I develop a distinct kind of blame, *oblique blame*, and unique kind of moral responsibility, *intervention-sensitive moral responsibility*. On this account, an agent is obliquely blamed in their capacity to intervene on a racist non-agential entity to motivate them to take action. This blame is oblique, rather than straightforward, because the agent is blamed at an angle, in their relationship to a non-agential entity. At the end of the chapter, I raise a problem for oblique blame: given our current epistemic resources, those who are in the best position to know that an entity is racist, those who experience racism, and are thus the most effective blamers, are often not trusted as credible knowers. This means that oblique blame often fails and accountability for racism is difficult to achieve. I call this the Knowledge Problem.

In Chapter 5, I argue that a collective shift in our epistemic resources is necessary for accountability for racism. I use the literature on norms to situate epistemic practices as governed by norms, showing that norms are an effective location for intervention related to this issue. I then review one overarching candidate norm and three specific candidate norms that organizing and activist groups with which I'm involved incubate through their practices. It is my hope that reviewing these strategies provides support to the idea that interventions into epistemic practices is one strategy for acting toward an anti-racist future. I end this chapter with some thoughts about how local norm shifts might be installed or spread into our dominant epistemic practices.

Finally, before I move onto some limitations of this project, I'd like to make a note about formatting conventions. Throughout the project I will use typical use/mention formatting conventions for discussing concepts, predicates, and properties. The conventions I will use are: caps for mentioning a concept, e.g. the concept RACIST is a primitive concept; single quotes for mentioning a word in English, e.g. the predicate 'racist' has six letters; and italics for the property to which the words refer, e.g. both individuals and norms can bear the property of *racist* such that 'racist' can be correctly predicated of them. In addition, I have followed the Chicago Manual of Style, 17th Edition guidelines for capitalization of ethnic and national groups (e.g. African American or Latina) and have also decided to capitalize People of Color and Black. When Touré (2012) discusses his editorial choice to capitalize Black, but not white, he emphasizes that Black

is an ethnicity in a way that white is not, pointing to the ways in which 'Black' has been used for the development of identity and solidarity in a way that 'white' has not been used. This editorial choice is common in works that focus on anti-racism.

1.5   Limitations

The first limitation has to do with epistemological limitations and my own identity as a white woman. Within anti-racist organizing and activist groups, both in person and online, there's a general guideline that white people shouldn't tell People of Color what is and is not racism. This is because People of Color, as those who experience racism, are in the best position to make this determination, especially as it relates to their own experiences. I try to follow this principle, and in particular, I make a point of avoiding telling people that they didn't experience racism when they think they did (in practice, this isn't very hard, since I also tend to think they experienced racism, but that doesn't matter much given the context). However, it may seem that this dissertation flies in the face of this of this principle. Specifically, in Chapter 3, I offer an account of how and why the predicate 'racist' should be applied. As is clear in this Introduction, my motivation for doing so is to push back against the idea that the concept of racism and the predicate 'racist' are hopelessly confused, bound to be so "lumpy" that we cannot arrive at a coherent and effective concept. I maintain that in interpersonal contexts, the above principle should be followed.

The second limitation is with respect to racial kind metaphysics. As will be discussed in Chapter 3, I've utilized a broadly social constructionist account of races throughout this dissertation. By this I mean to reject biological realism about race and accept a social realism view about race. I have also chosen to use the turn of phrase "racialized as…" at points throughout the dissertation in order to highlight the ways in which race is continually constructed, ascribed with new social meanings as we learn more and as social imaginaries shift. In taking up these strategies, I utilize folk conceptions of race, including the racial terms Black, Latinx, Asian, Indigenous, and white. However, this strategy isn't without its limitations.

First, though this is not entailed by a social constructionist position, utilizing racial kind terms may work to reify and cement racialism (Appiah 1992, 13-14), or the view that there are a small set of racial groups with inheritable traits and tendencies. At the very least, the use of racialized terms works against what Spencer (2017, 11-12) calls "radical racial pluralism," or the view that there are no concrete answers to what races are or to whether or not races are real even

in this social sense. Second, utilizing folk conceptions of race may also erase the experiences of people who reject racialization or serve to bolster the idea that individuals among racialized groups have similar experiences or a shared history. For example, in describing her terminological choices, Dunbar-Ortiz (2014, xiii) writes, "I also wanted to set aside the rhetoric of race, not because race and racism are unimportant but to emphasize that Native peoples were colonized and deposed of their territories as distinct peoples—hundreds of nations—not as a racial or ethnic group." In utilizing broad racialized terminology in this project, I am unable to capture any of the unique stories or histories of these groups. Finally, many empirical studies use ethnic group names like "African American" and "Hispanic." When discussing these studies, I use the terms that they use to avoid confusion, but this leaves many metaphysical questions on the table.

The third limitation is that this project lacks an intersectional analysis in that I focus almost exclusively on race. In the Combahee River Collective Statement (1977) the authors emphasize the ways in which the white feminist movement and the Black Power movement had failed to include them in various ways as Black women. This caused them to develop the collective as an organization that understood the interlocking systems of oppression that Black women face, as well as other people who experience multiple forms of oppression. They write, "We also often find it difficult to separate race from class from sex oppression because in our lives they are most often experienced simultaneously." With this insight in mind, it is likely only superficially that I have been able to separate out racism from these other systems of oppression. In addition, this choice leaves out conceptions of racism/sexism such as misogynoir. Thus, I do not claim to give a comprehensive account of the sufferings of people's lives, but to illuminate one feature. I've chosen this strategy because white people so often have a hard time zeroing in on race, keeping their attention on anti-racism, because they desperately do not want to talk about it (DiAngelo 2018, Oluo 2018).

I've tried to keep these limitations in mind as I have constructed and developed this project, but there are no doubt times when these limitations do obscure my thinking and the conclusions to which I have come. It is my hope that the work in this dissertation can play a role in making our conversations around race more successful in ways that lead toward anti-racist commitments and action.

CHAPTER 2.   LEVELS OF PLURALISM AND NESTED QUESTIONS, OR WHY WE SHOULD BE ENTITY TYPE PLURALISTS WITH RESPECT TO RACISM

2.1   An Enduring Debate

The large goal of this dissertation project is to (1) improve clarity around the concept of RACISM, particularly with respect to accounts of racism and the predicate 'racist', (2) in order to preserve the moral function of the predicate 'racist' (3) so that we can hold each other accountable. In this chapter, I'll focus on (1), homing in on the enduring debate between individualist and structuralist accounts of racism within social and critical race philosophy. This chapter is an extension of the recent philosophical work that hopes to both clarify and potentially deflate the deep dichotomy between the individualist and structuralist projects. Specifically, I hope to extend the work done by Madva (2016) and by Kelly and myself (2018). In his "Anti-Anti-Individualism" paper, Madva (2016) argues that structuralists do not appreciate the personal commitments and transformations needed for sustainable social change. Kelly and I intervene on the debate between individualists and structuralists by arguing that norms mediate between the two levels of explanation, and thus, the projects are not neatly separable from one another.

These projects have added to the growing suspicion that there's something wrong with the individualist versus structuralist debate. The tools developed throughout this chapter will allow us to diagnose these issues and think more clearly about what insights to bring forward and what commitments need to be left behind. The debate persists due to a pervasive metaphysical misunderstanding, and the new frameworks developed in this chapter will help absolve these confusions. Specifically, I'll argue that individualists and structuralists are both wrong with respect to the metaphysical commitments taken on within their views. I'll show that both individualists and structuralists are entity type monists who mistakenly restrict applications of the predicate 'racist' to particular kinds of entities and, in doing so, make mistakes that are damaging to our understanding of racism and to the anti-racist project (particularly with respect to the practice of accountability). Before this payoff can be delivered, I need to develop tools for better understanding the debate. Here's how this chapter is set-up:

In Section 2.2 of the chapter, I'll present several existing taxonomies that describe and differentiate between metaphysical social kind projects and accounts of racism in particular. It is my hope that doing so will allow us to better see locations of agreement between scholars, as well as the places where different accounts diverge. First, I'll discuss Haslanger's (2012a) taxonomy for metaphysical projects[7] to lay the groundwork for identifying what philosophical accounts of racism are trying to do. Second, I'll discuss Harris' (1998) taxonomy for accounts of racism. Third, I'll describe some basic features of individualism and structuralism, the broad families of views that accounts of racism are often grouped into and highlight Madva's (2016) distinction between backwards- and forward-looking questions.

In Section 2.3, I'll put forth an additional frame that I argue is needed to understand similarities and differences between accounts of racism. I'll call this frame *levels of pluralism*. I'll argue that this frame allows us to better see the differences and points of disagreement between accounts. In particular, it will give us a more fine-grained way of analyzing the differences between individualist and structuralist accounts of racism.

In Section 2.4, I'll offer my account of nested focus questions, which fills the *levels of pluralism* frame. Here, I identify four sets of focus questions one might ask in developing an account of racism with corresponding sets of motivations. The four types of questions are: questions about the moral status of an entity or outcome, questions about what kind of entity was involved, questions about the causal story, and questions about what kind of interventions would best address what has occurred. Though none of these sets of questions are more important than the other, I'll argue that the questions must be *nested*. Just like baskets of four different sizes must be placed in a particular order to *nest* inside of one another neatly, so too must these questions about racism. I'll argue that the questions are in a dependence relationship with one another in that answers to the "biggest" questions depend on or are affected by answers to the "smaller" questions. Because of this dependence relation, answers to "bigger" questions may obscure our thinking about the "smaller" questions. For example, we may think that an inability to intervene effectively on something must mean that it cannot be racist or that it wouldn't be appropriate to predicate

---

[7] I take it that Haslanger's (2012) driving questions have to do with the metaphysics of social kinds. However, these projects importantly bring naming questions (once we get the metaphysics right, how should we talk?) and epistemic questions to fore. See Barnes (2019) for a discussion of the relationship between these kinds of projects.

'racist' of it. This is allowing questions of *intervention* to influence questions of *moral status* or *evaluation*.

In Section 2.5, I'll review individualist and structuralists accounts that restrict applications of the predicate 'racist' for particular kinds of entities. I'll call these accounts entity type monist views. I'll show that there exist true individualists and structuralists who argue (or implicitly assume) that some entities (individuals or structures) bare the property *racist* in some special, fundamental way.[8] Though most theorists do not have an explicit theory for when and how the predicate 'racist' should be applied, they do signal that they think some things are *really* racist while other things are *unimportantly* or *merely consequentially* racist. This is often signaled by phrases like "x kind of thing is of primary moral concern." That there are actually individualist and structuralist views matters, because otherwise we have sustained a debate that isn't grounded in actual disagreement.

In Section 2.6, I'll show how the new frame of *levels of pluralism* and my account of nested questions explains why theorists end up as individualists or structuralists. Specifically, I'll demonstrate that a theorist's primary motivation influences the focus question that they hone in on. Further, in prioritizing answering a particular focus question, answers to other focus questions are obscured due to the dependence relationships between the focus questions. Entity monist views come about because "big" questions are answered without recognizing the dependence relationship with "small" questions. In other words, individualists and structuralists "unnest" the four focus questions. The result of this discussion is my argument for why the dependence relationship between the questions should be attended to and why we should be entity type pluralists with respect to racism. Here's a decision tree to keep in mind throughout this chapter:

---

[8] I do this because one approach to the debate between individualists and structuralists is just to say that the debate had been overemphasized and that no theorists is *actually* one or the other such that have views that are *at odds* with the other view.

Figure 2 Decision Tree for Levels of Pluralism, Individualists versus Structuralists

I'll conclude in Section 2.7 with some concerns Blum (2014) will have with the view I've presented and preview some of the questions that will be answered in Chapter 3.

2.2   Carving the Accounts

Harris (1999, 21) has argued that any good theory of racism will have an account of "whether or not races exist, what attitude we should have toward races, how to explain the influence of objective or socially constructed racial groups, how to define racism, and some indication of how to judge controversial cases." Accounts of racism have a big job to do, and thus they are understandably complex. However, not every theorist is able to meet all of these *desiderata*, and they make choices about what features of a full account to which they will direct their attention (just like I've decided to hone in on the property *racist* and the predication of 'racist'). These choices influence the kinds of accounts they give. To understand the different projects that different accounts of racism engage in, scholars have offered taxonomies for making the methods and aims of projects clear. This allows us to see when, if, and why scholars disagree or are offering compatible accounts with different foci. In this section, I'll discuss three taxonomies for understanding the choices people make, the way in which this influences their focus, and the impact it has on their accounts of racism.

2.2.1 Haslanger's Taxonomy: Conceptual, Descriptive, and Analytical

Most accounts of racism can be understood as theories that attempt to answer the question: "What is racism?" In answering this question, many more questions will likely have to be answered. In discussing theories of this type—the type that try to answer "What is X?"—Haslanger (2012a) provides a tripartite taxonomy of projects in which one might engage and the various methodological commitments that follow from each project.[9] On her taxonomy, one might be engaged in a *conceptual* project, a *descriptive* project, or an *analytical* project (223-224).

According to Haslanger's taxonomy, *conceptual* projects seek an articulation of ordinary concepts using either reflective equilibrium or conceptual analysis. Because conceptual projects have to do with the way concepts are deployed in ordinary language, these accounts will be sensitive to the ways in which terms are actually used. *Descriptive* projects, in contrast, focus on the extension of concepts, perhaps improving (our epistemic position with respect to) concepts by learning more about the natural or social phenomenon with which we are concerned. With respect to race, these kinds of projects help us learn whether our use of race terms tracks a (social) kind. If they don't track anything meaningful, we might try and shift our (epistemic position with respect to) concepts so they track the social realities we mean to track.[10] And, finally, *analytical* projects consider the pragmatics of concept use and uses these practical concerns to offer (often revisionary) concepts.[11] The analytical approach is constrained by the natural or social phenomenon in which one wants to talk about, but non-metaphysical, non-epistemic goals enter into these kinds of metaphysical/epistemic projects. The constraint is in place because highly technical terms aren't

---

[9] In this (2012a) article, Haslanger focuses on "What is X?" questions in which the X is a social kind. This may be importantly different than questions in which the X is some other kind of social phenomena such as a complex social process. For the purposes of this chapter, I assume that Haslanger's taxonomy of projects can be applied to accounts where the X is "race" and when the X is "racism."

[10] Descriptive projects are going to be muddled with commitments concerning concepts, reference, and meaning. Let's use the concept FROG as an example. Imagine a layperson, Fred, who hasn't studied much about frogs, but wants to learn more. When Fred uses the term 'frog' to pick out something in the world, Fred has a bunch of incorrect information about frogs in mind. His epistemic position with respect to frogs is very poor. One could describe the situation this way: Fred's (epistemic) folk concept FROG doesn't pick out any actual frogs because the extension of the term doesn't cover actual frogs only what Fred takes to be frogs. Or, with a different set of methodological commitments, one might say: when Fred uses the term 'frog' he successfully expresses the concept FROG that tracks frogs in the word, but that Fred is just very mistaken about what frogs are like. In the first case, to improve Fred's epistemic position, we'll want to shift his folk concept FROG. In the second case, we need to teach Fred more about the extension of 'frog'. In either case, we'll look to our frog experts to get Fred up to speed. We'll want to improve Fred's epistemic position with respect to FROG regardless of our chosen metaphysics of reference and meaning.

[11] Analytical projects share some features with explicative definitions in that they improve upon existing concepts. Gupta (2015, 1.5) says that explicative definitions "may be offered as a 'good thing to mean' by the term in a specific context for a particular purpose."

going to do very well with respect to the pragmatics of concept use and deployment. Haslanger takes herself to be engaging in analytical projects with respect to our concepts of gender and race. Thus, Haslanger (2012a, 221) asks not just "What is X?" but also "What do we want to them to be?"

Two elaborations of Haslanger's taxonomy will also be helpful. In his work on the social construction of human kinds, Mallon (2016) identifies his project as falling under *explanation-driven* metaphysics, contrasting Haslanger's (2012a) analytical project as *justice-driven* metaphysics. Mallon (2016, 209) writes, "…our accounts differ fundamentally in the essential reference Haslanger's account makes to the normative concepts of 'oppression' and 'subordination'…This is no mistake, for, by Haslanger's lights, it is these normative concepts that mark the domain of relevance for progressive social theory." Importantly for Mallon, however, is that explanation-driven metaphysics can still illuminate normative projects; they just don't use normative concepts to make metaphysical claims. Though Mallon's own label 'explanation-driven' is more specific, Mallon's project largely falls under what Haslanger calls *descriptive* projects.

Barnes (2019) also splits up the project pie, separating analyses into *descriptive* and *ameliorative* projects. As with Mallon, she only has two categories. She seems to bring Haslanger's *conceptual* and *descriptive* projects under one umbrella, using *descriptive* to define them both. She suggests that on *descriptive* accounts, the extension of terms is fixed by their contemporary use (rather than by experts who are taking a close look at social/empirical science). Though I think we lose some important conceptual nuance with this grouping, Barnes has a very important clarification to offer both kinds of projects namely the possible separation between the metaphysical accounts of (social) kinds and the suggested use of terms with respect to kinds. She argues that our metaphysical accounts (of gender) should constrain and guide our use of (gender) terms, but that metaphysical accounts need not give the necessary and sufficient conditions for genuine category membership and term deployment. Thus, on her account, metaphysical accounts can come apart from how we improve our use of language.

Haslanger's (2012a) taxonomy, alongside these expansions and clarifications, moves us forward in understanding exactly what is going on in philosophical accounts of racism. Haslanger means for this taxonomy to help us out with respect to all metaphysical projects that ask, "What is X?" Further, she focuses on social kinds rather than concepts like sexism or racism. Thus, it will help us further to add a taxonomy specific to racism.

2.2.2   Harris' Taxonomy: Descriptive and Explanatory

In contrast to Haslanger (2012a), Harris' taxonomy is usefully limited to accounts of racism. So, it is sensitive to specific features of accounts of racism. The primary chasm on Harris' taxonomy is between descriptive accounts and explanatory accounts of racism. As will become clear, Harris (1998, 2018) thinks that explanatory accounts have problems that descriptive accounts do not. For now, I'll focus on the kinds of accounts that fall into the respective categories, starting with those accounts of racism that are explanatory.

Harris (1998) identifies three competing explanatory accounts of racism: hierarchical-belief, institutional, and logic-based accounts. The first set of accounts theorize that racism is primarily involved with "ranking races in a hierarchical fashion of irredeemable or relatively stable inferior/superior kinds" (Harris 1998, 225). On these accounts, *beliefs* about races are the primary location for explanation with respect to racism.[12] Institutional accounts, on the other hand, do not take beliefs about racism to be causally significant. Rather, these accounts (1998, 225) are "dependent on descriptions of outcomes for following usually color-blind rules and regulations." The locus of explanation on these accounts is more varied, including laws, policies, and ideologies like white supremacy.[13]

The third set of views are logic-based accounts in which "inconsistencies, contradictions, attitudes of self-deception, deceit, living in bad faith, living in avoidance of reasoned judgment, etc. are the primary notions used to explain, describe, and condemn racism…that there is something wrong with *the way one thinks* that is the primary focus of discussion" (emphasis mine, 225).[14] Thus, on these accounts, individual psychologies are the location of explanation, though individual psychologies can be said to reflect structural or ideological features.

In this 1998 paper, Harris separates theories of racism into these three categories, noting that theories in each category attempt to provide an explanation for racism or the outcomes of racism. In his 2018 paper, Harris (2018, 40) categorizes all theories that "provide a neat picture of

---

[12] Appiah (1992) and Blum (2007) are notable examples.

[13] Harris explicitly includes Carmichael (Ture) and Hamilton (1967), Omi and Winant (1986, 2015), and Outlaw (1996). Since Haslanger (2018) theorizes racism as a set of social practices that constrain the available options along racial lines, I think she belongs here as well. Bonilla-Silva's (2003) "racism without racists" account can come under this label as well.

[14] Harris lists Cashmore (1987), Delacampagne (1983), Gordon (1993, 1999), Wieviorska (1994), Kovel (1970), and Allport (1954). Due to his theory of the racial contract as a mistaken set of cognitions, Mills (1997) seems to be a good addition here, too.

its [racism's] causes and simultaneously provide reasons, consonant with its explanation, for its moral wrongness" as "racism as a logic" accounts. Thus, though each of these three kinds of accounts differ in specific features, they can all be grouped as *explanatory* accounts.

Harris goes onto argue that explanatory accounts are insufficient, especially with respect to anomalies. Consequently, we need another branch of the taxonomy, so Harris (2018) offers his own *descriptive*, actuarial account of racism, which he calls *necro-being*. Descriptive accounts, as opposed to explanatory accounts do not offer unified causal accounts, and they do not require that the moral wrongs of racism are connected with causes (unified or not). Descriptive accounts begin with just that: description. Thus, Harris' descriptive account begins with identifying and describing the suffering, misery, and premature death of people of color around the world.[15]

He offers this descriptive account against the backdrop of the many *explanatory* accounts of racism described above. In focusing on providing unified explanations for racism, most other accounts, Harris (2018, 40) argues, are subject to anomalies. In other words, there are things we think count as racism that fall outside of the purview of the theory. The accounts then either provide special accounts of how to include the anomalies or allow the anomaly to be an unfortunate upshot of their view. In discussing a wide variety of cases—from the Rwandan genocide to the rape of Rohingya women—Harris claims that explanations of racial miseries vary greatly across societies and the particularities of each societies' histories and local circumstances. He (2018, 44) writes, "I could not see the intricacies of the miseries for themselves as themselves if I presume a logical social structure and a well-defined derivation manual that maps particular circumstances of events to the structure." In other words, explanatory, racism as logic, accounts of racism, which provide a unified explanation for racism fail to take into account the vast array of subtleties of individual miseries that lead to premature death. Further details of the account of racism Harris (2018, 40) defends will be discussed in Chapter 3, but for now, it is helpful to note that a *descriptive* account of racism "helps us to see a vast range of miseries in ways that explanations fail to convey." Thus, we see the important difference between *descriptive* and *explanatory* accounts.[16]

---

[15] Around the world here is important. One reason Harris (2018) argues that explanatory accounts of racism fail is because they are context specific. In other words, the explanations of racism depend too fully on specific cultural features, and thus, cannot be used to explain racism outside of a specific context. This is problematic because many accounts of racism claim (implicitly or explicitly) to be *unified* accounts of racism, to explain all instances of racism with one causal story.

[16] It is important to note here that Harris' taxonomy is importantyl normative in ways that Haslanger or Mallon's are not. In contrast, both Haslanger and Mallon say that one might engage in the different projects for different reasons

One note about how Harris' taxonomy maps onto to Haslanger's. Harris (2018) identifies his project as *descriptive*, but he doesn't invoke Haslanger's (2012a) sense of descriptive, though it could be compatible. I take Harris' focus to be on describing the morally relevant features and outcomes of racism, rather than offering a neat explanatory framework that identifies a unified casual story for such outcomes. So, if we take Mallon's project as descriptive on Haslanger's picture and Harris' project as descriptive on the same picture, then there will be extremely important differences within the *descriptive* projects category since Mallon's account is specifically explanatory. Thus, it maybe that Harris' project is a *descriptive-descriptive* account, rather than an *explanatory-descriptive* account.

No matter how Harris' (2018) account sits within Haslanger's taxonomy, Harris (1999) also clearly has a revisionist project in mind. In discussing what accounts of racism can do, Harris (1999, 17) writes, "A concept, as distinct from a definition, can tell us explicitly what a word *should* mean." This means his view is not a *conceptual* project on Haslanger's account and sets him distinctly apart from folks like Blum (2002) who argue that our accounts of racism should (or must) reflect historical usage of the term. Harris' taxonomy, which boils down to descriptive and explanatory projects is extremely useful for helping us understand what accounts of racism are up to. Now for a more popular, but perhaps markedly less helpful, divide.

2.2.3   Individualism and Structuralism

The debate between individualists and structuralists was revived by Haslanger's (2015) paper critiquing implicit bias research for its "individualism." Alongside this critique, Haslanger (2015, 1) claimed that implicit bias *must be situated* within a structural explanation because "changing structures is often a precondition for changing patterns of thought and action." Here, Haslanger identifies a longstanding debate within social theory: is the best location for description, evaluation, explanation, and amelioration[17] at the level of the individual or at the level of structure? Depending on the answer to this question, accounts of racism (or oppression in general) get placed into two broad families of views: individualism or structuralism. If one thinks that individuals (and

---

and all of these types of projects may uncover important things about our social and lived world. Harris, on the other hand, is pointing out a difference in strategy that leads to mistakes.

[17] Actually, the debate is far less clear than this. In Section IV, I'll argue that thinking about debates with these four motivations (and corollary sets of questions) will help us see the points of agreement and disagreement between accounts.

what's in their heads and hearts) are the best starting points, then one is an individualist. If one thinks that practices, policies, and laws are the best place to begin, then one is a structuralist.

Though a lot seems to ride on this divide, these two families aren't very well defined. But, perhaps we are making progress. For example, Haslanger's (2015) paper spurred an "anti-anti-individualist" response by Madva (2016), which both further cemented the debate between individualists and structuralists and raised the question about whether one needs to choose between the two camps. Then myself and Kelly (2018) muddied the deep divide between the two kinds of accounts by adding a discussion of norms into the debate (norms are especially good muddying agents because norms are group-level regularities that are also encoded in individual minds, meaning norms mediate between the two "levels" of explanation). Thus, the debate isn't as clear cut as it seems on the surface.

To clear things up a bit, Madva (2016) distinguished between backward-looking and forward-looking questions within accounts. The backward-looking aspect of a given theory identifies the causes of moral wrongs, where the forward-looking poses locations for interventions and solutions. The backward- and forward-looking accounts are often similar for individual theorists, i.e. they identify the same entity as both the cause and solution of a given social ill (see Anderson (2010)), but they need not be. For example, one may think that the root cause of inequity is unjust social structures, but that the answer is changing individual hearts and minds (or vice versa, etc.). With the introduction of this distinction between backward- and forward-looking accounts, Madva begins to make space for precise disagreement within each broad family and provides a bridge between the two families. With this new lens, it turns out that one might be an individualist with respect to explanation but a structuralist with respect to intervention (or vice versa, etc.).

At this point, we've got three taxonomies that we can employ to better understand what's going on within particular accounts of racism and the differences between separate accounts of racism. Though each taxonomy brings something important to the table, they are also articulated with different concerns and aren't optimally compatible. Philosophers sometimes do just throw their hands up at the debate (see the discussion of Madva's (2016) paper on the Brains Blog), but we haven't really moved forward in understanding the debate and what it would take to absolve it (what if we don't want to be individualists or structuralists?). In what follows, I provide a new taxonomy, a new frame for understanding accounts of racism.

2.3   Levels of Pluralism

In the next two sections I will introduce a new taxonomy, a new frame, for understanding accounts of racism and the debate between individualists and structuralists. It is my hope that this clears the ground for showing that both the individualists and structuralists are mistaken and are wrong to limit the kinds of things they appraise as racist or not. I call this new frame *levels of pluralism*. An assumption of this frame is that we will be able to better compare views and argue between them (or against them) when we see whether a particular account is pluralist or monist with respect to different aspects of racism.[18] I've identified four levels of pluralism. I've identified these levels because they capture the questions to which the answers matter for accounts of racism. Positions on each of these levels are defining features of views.[19] Whether explicit in theories of racism or not, accounts of racism take a position on each of the four levels. In other words, they are either monist or pluralist with respect to each level. In addition, the last three levels together make up explanatory pluralism/monism. In order for explanatory pluralism to be maintained, a theory must be pluralist with respect to all three levels. I've articulated the below table through the pluralist lens, but an account can just as well be monist with respect to each level.

---

[18] A quick note about terminology: throughout this chapter, I will be using "monist" as an antonym to "pluralist." When it comes to identifying the kinds or types of entities that can bare the property *racist* (entity type monism/pluralism), this use is straightforward since 'monism' and 'pluralism' are typically used to describe metaphysical claims. However, when it comes to describing views as causally monist or pluralist, I stray from a strict metaphysical sense of the term and instead draw a contrast between "singular" and "plural" or "reductive" and "non-reductive." My primary motivation for this use is to keep the concepts clean, while still pulling some descriptive force from the terms. But, it is important to keep in mind that, strictly speaking, views that I identify as morally monist, for example, can be metaphysically pluralist in that the moral property shows up in many different ways in different entities. These views are monist in the sense that they identity the wrongs of racism in a *singular* family of features or offer *reductive* accounts of various moral wrongs.

[19] This doesn't mean that more levels of pluralism couldn't be identified. However, selecting levels at decision points that make a difference for the view is the best strategy for diagnosing differences between views and the metaphysical misunderstandings between individualists and structuralists.

Table 1 Levels of Pluralism

Levels of Pluralism			
Moral Pluralism	**Entity Type Pluralism**	**Causal Pluralism**	**Interventions Pluralism**
Pluralism about the wrongs of racism	Pluralism about the primary entities involved with racism	Pluralism about the causes of racism	Pluralism about the appropriate interventions with respect to racism
Moral Pluralism	Explanatory Pluralism		

Thinking about pluralism with respect to racism isn't new. Motivated by trying to capture the wide variety of experiences of racism across the world, Appiah (1990) articulated an account of "racisms." On this view, there isn't one thing that's racist, but many things. In addition, racism may take many forms such that all instances of racism do not share in the same moral profile. What makes racism wrong can be different in different cases. This is a form of moral pluralism. For moral pluralists, there are a wide-variety of morally relevant features that can make something that is *racist* wrong. Moral pluralism is also the motivation for Blum's (2002, 2014) insistence that we identify many different racial ills that are wrong but are not racist. Moral pluralism is perhaps the most common explicitly pluralistic move within the literature articulating accounts of racism.

However, this is not the only kind of pluralist one could be. And, just because a theory isn't explicit about its commitment to pluralism or monism, doesn't mean that tacit commitments aren't present. So, in order to increase our understanding of different views, I put forth three new kinds of pluralism to be explored: entity type, causal, and interventions pluralism.

Entity type pluralists with respect to racism are committed to the view that there are a variety of things to which the predicate 'racist' can be applied and that there is no core set of things or type of entities that is the fundamental barer of the moral wrongs of racism. Throughout the dissertation, I will use 'entity' very broadly to identify locations of analysis relevant to racism (e.g. mental states, actions, social patterns, social heuristics and epistemic resources, norms, policies, campaigns, laws, and institutions). Those that do think there are core set of entities involved in racism are entity type monists. As will be discussed in Section 2.5, individualists and structuralists with respect to racism are entity type monists because they think that the normative core of what is wrong with racism lies in either individuals or structures, but not in both.

Causal pluralists believe that there are a variety of causes of racist outcomes and that there is no logical structure that ties these multiple causes together into a single explanation for such

outcomes. For example, Harris (2018, 1) calls racism a "polymorphous agent of death," meaning that racism takes many forms with a variety of sustaining causes. This expresses his commitment to causal pluralism. Causal monists, on the other hand, are committed to the view that all genuine instances of racism are caused by a particular mechanism. These views are discussed above as explanatory accounts in the descriptive versus explanatory taxonomy offered by Harris (1998). Not every causal monist account identifies the same mechanism. The mechanisms or causes identified from account to account vary greatly—beliefs, institutions, flaws of logic or epistemology—but what causal monists share is that they think all instances of racism are caused by the same process. For example, Gordon (1999, 75-76) argues that anti-Black racism arises out of a form of Sartrean bad faith that evades responsibility and is stubbornly resistant to counter-evidence. In identifying one cause of racism—bad faith—we get causal monism (and, subsequently, entity type monism).

Finally, intervention pluralists think that there are a variety of interventions that might be successful in eliminating or limiting racist entities and racism outcomes. Some interventions might be interventions on individuals, others might be on norms, and others might be on policies and laws. A commitment to interventions pluralism calls the individualist and structuralist divide into question. For example, Madva's (2016) paper arguing against "structural prioritizers" is focused on the ways in which structuralists (who are often causal monists) are intervention monists, as well. For example, he cites Anderson's (2010) view that because segregation is the cause of morally bad outcomes (a causal monist view), then integration must be the appropriate intervention. Madva points out that many structuralists identify the cause and appropriate intervention site as the same entity. Thus, as discussed above, Madva suggests that we separate "backwards-looking questions" (causal questions) from "forward-looking questions (interventions questions).

However, even with this distinction in hand, there are times when interventions pluralism is used to argue against structuralist views as a whole. This leads to confusion because interventions are not the only kind of commitments within accounts of racism. A structuralist might say in response, "Yeah, I understand that individuals must be committed to social change in order for sustained institutional change to occur, but I still want to be a structuralist." The levels of pluralism account I've articulated here provides a concrete way forward because, rather than using the ambiguous term 'structuralist', one can say, "I'm a causal monist and an interventions pluralist."

In this example, we can see why it is useful to further divide up the conceptual space so that we can precisely target our objections.

As mentioned above, entity type, causal, and interventions pluralism can be grouped together under the heading of explanatory pluralism. In other words, if one is a pluralist with respect to all three, one is an explanatory pluralist. This is because the latter three questions attempt to explain how or why racist outcomes take place. This is meant to track with Harris' distinction between descriptive and explanatory accounts of racism. Within Harris' taxonomy, explanatory accounts of racism are explanatory accounts precisely because they are explanatory monist accounts. In other words, explanatory accounts of racism aren't explanatory because they give explanations in general but because they attempt to give unified explanations or are explanatorily monist. In contrast, descriptive accounts are, in virtue of their strategy and metaphysical commitments, explanatorily pluralistic.

As is indicated by the title of this chapter, I think something special rides on entity type pluralism. Specifically, the debate between individualists and structuralists relies on the rejection of entity type pluralism and the acceptance of entity type monism. This is something shared by the views. Accepting entity type pluralism is a way to reject the dichotomy, reject having to develop a view that lays neatly on either side, and allow for new and refreshed debates to enter into our discourse. In order to understand why this is the case, I need to further fill out the levels of pluralism frame. In the next section, I will show that commitments to pluralism or monism on each of the levels in this framework are motivated by answers to corresponding focus questions. Along with each set of focus questions are corresponding motivations for engaging in philosophical work on racism. These focus questions and motivations will fill in our understanding of the levels of pluralism frame.

2.4   Nested Focus Questions

2.4.1   Focus Questions and Motivations

The strategy of delimiting questions is not unique to me. In a landmark paper, Tinbergen (1963) offers four questions that can be asked about animal behavior that correspond to four different kinds of explanations of why an animal has a particular feature or ability. In this paper, he presents the questions as separate, but equally important, depending on the interests of the

questioner. In fact, if one was to get a full view of an organism, one would have to have some sort of answer for each of the four questions. In this section, I will articulate a set of four questions that are answered within accounts of racism. Answers to each of these questions will be important because, as discussed by Harris (1999), accounts of racism have a lot of work to do. Parsing the questions into categories helps us clarify the kinds of claims being made within accounts of racism. The four sets of focus questions align with the four levels of pluralism described above such that the answers to the questions are what makes someone a pluralist or a monist with respect to a particular level.

Here are (some of) the kinds of questions that might be answered for each set of focus questions:

1. Moral Status: Is this (happening in the world) wrong? If so, what makes it wrong? Is this wrong racist?
2. Entity-Type: What kind of thing is involved in the wrong? Does the entity involved with the wrong have volition or agency? If so, can it/they do something about stopping the wrong? If not, what can we (entities with agency) do about it?
3. Causal: How did the wrong happen? What caused it? What mechanisms support the continuation of the wrong? What's the history of this wrong? Is the wrong systematic or a one-off wrong?
4. Intervention-Focused: How do we keep this wrong from happening again? How are the answers to the first three questions useful in helping us identify appropriate interventions? Do we need any additional information (i.e. information about the function of mental states or processes or about how to change a formal law or procedure)?

It may be the case that any robust theory of racism will answer in detail all of these questions; however, many people start or focus on one set. The set with which they begin will depend on their motivation for articulating the account. I've found there to be four primary motivations: description, evaluation, mechanism identification, and amelioration (mechanism identification and amelioration are often tied very closely together). Harris (2018) focuses on **description** of the misery and suffering of racism. Garcia (1996) focuses on **evaluating** others as racist or not. Kovel (1970), Mills (1997), Gordon (1999), Shelby (2003), Anderson (2010),[20] and Haslanger (2015), focus on **mechanism identification** and **amelioration**. I've separated out motivations of mechanism identification and amelioration here to reflect Madva's (2016) contribution of backwards- and forward-looking questions. Many folks who give explanatory accounts are

---

[20] Anderson (2011) is an interesting outlier here in that she is most clearly a thoroughgoing structuralist with respect to racial injustice but that she has a very individualistic conception of racism. Thus, I will continue to analyze her work in both categories.

ultimately motivated by amelioration but focus on mechanism identification because they think identifying good interventions will require knowing more about the mechanism.

Table 2 Nested Questions and Motivations

	NESTED QUESTIONS			
MOTIVATION	Description	Evaluation	Mechanism Identification	Amelioration
FOCUS QUESTION	Moral Status	Entity Type	Causal	Interventions

2.4.2 Nested Focus Questions

As Tinbergen (1963) articulates the four questions within the field of ethology, each set of motivations and questions is distinct from the other. In other words, it isn't the case that answers to one set of questions have the ability to influence answers to other questions. This is where my account of questions departs from Tinbergen's strategy. Though any full account of racism will answer each question at some point, in this section I will argue that the focus questions are *nested*. In other words, I'll argue that the questions are in a dependence relationship with one another such that answers to some questions have the ability to influence answers to other focus questions. This dependence relationship has the ability to cause problems in our accounts of racism, so an understanding of the dependence relationship is integral to avoiding mistakes.

For the purposes of the *nested* metaphor, I'll be conceptualizing *moral status* questions as the "smallest" questions, *intervention-focused* questions being the "biggest" questions, with *entity type* and *causal* questions in the middle. None of these sets of questions are more important than the other, but an appreciation of the dependence relationship is necessary. Just like baskets of four different sizes must be placed in a particular order to *nest* inside of one another neatly, so too must these questions about racism. The questions *nest* together such that "smaller" questions are inside of the answers to "bigger" questions and, without full appreciation of this feature, answers to "bigger" questions can obscure our thinking about "smaller" questions in a way that leads to mistakes.

As Harris (2018) argues, explanatory accounts are unable to explain or account for anomalies because the accounts couple together the causes of racism with the wrongness of racism. In these cases, the source or cause determines that a wrong is of the *racist* type. He (2018, 62) writes, "The miseries suffered by victims of racisms happen even if there is no neat causal chain

from individual invidious behavior to the misery of an individual or the workings of institutional rules to the misery of a whole population." Another way of putting this is that theorists allow their answers to entity type, causal, or intervention-focused questions to constrain their answers to questions of moral status. Separating questions of moral status from questions of entity type, causal, and interventions-focused and nesting these questions in the right order avoids these kinds of mistake.

On the view I'm offering here, answers to the questions of moral status will give us our descriptive (as opposed to explanatory) account of racism. In other words, it will describe the morally relevant features and identify them as having the property *racist*. Answering this question, however, won't give us a story about the kind of entity the racist thing is, how it came to be racist, or what we should do about it. In other words, the answers to moral status questions don't offer an explanation of how the world came to be that way. Though we will need to answer the next questions in order to know what to do next, in answering questions of moral status, we've already accomplished something very important. And significantly, we haven't inappropriately constrained our answers with our explanations. However, it isn't the case that answers to questions of moral status, the descriptive portion of our theory, won't appeal to any patterns at all. Rather, it won't appeal to *causal* patterns or simple determinants. Harris' (2018) view, for example, identifies patterns of premature death and misery that amounts to living death as the constitutive moral profile of racism. We know racism is present when race-based premature death is present.

Answers to moral status questions are the core of accounts of racism; they are in the smallest basket, representing the normative core. When these questions are asked without interference from the larger questions, there are no constraints on the kinds or types of entities that can be racist. Each entity found to be racist is a fundamental bearer of the predicate 'racist.' This is because we haven't allowed answers to entity type or causal questions to influence our questions of moral status. Focusing on this normative core rather than on explanatory frames means that each entity that has the property *racist* is not explained by or constitutive of another entity's having the property. Note that the having of a property at time t is independent from the cause of an entity having a property at time t. Thus, one entity may cause another entity to gain the property without

being constitutive of its having the property.[21] These causal forces are irrelevant to the entity's evaluation as racist or not.

On this account of nested questions, answering *entity type*, *casual*, or *intervention-focused* questions before answering questions about *moral status* causes theorists to exclude unnecessary entities from their accounts of the primary wrongs of racism (which makes them individualists or structuralists, see Section 2.5). As will be argued for more fully in Chapter 3, answers to questions about moral status will overlap greatly with the application conditions for the predicate 'racist'. Thus, another way of thinking about my argument is that individualists and structuralists both have inaccurate application conditions for the predicate racist. Worrying about the kinds of entities under appraisal, the causes of or unified explanations of racist outcomes, or identifying the most apt interventions causes theorists to unnecessarily and mistakenly limit their appraisals of 'racist' to certain kinds of entities. Which entities are missed will depend on the focus of the theory; individualists miss informal and formal structures (see Davidson and Kelly (2018) for a discussion of informal and formal structures) and structuralists miss individuals.

Now that we've got the conceptual framework for understanding primary motivations and nested focus questions onto the table, we can put them together to understand decisions within the levels of pluralism frame. When the motivations, focus questions, and levels of pluralism are taken together, they give us a better framework for understanding how and why different accounts of racism end up having the commitments to monism or pluralism that they end up having.

---

[21] These distinctions will be important in discussions of reducibility. For example, Garcia (2011) argues that institutional racism is both morally and explanatorily reducible to the racism of individuals (11).

Figure 3 Decision Trees for Levels of Pluralism

Notice that on my view of nested questions, once monism is introduced into an explanatory level (entity type, causal, or interventions), then one will become a monist on all subsequent levels. [22] Introducing monism into one's explanation, as is often done when mechanism identification is the primary focus of an account, pervades all other aspects of the explanation. As will be important for the next chapter, moral monism (a descriptive feature) doesn't have this effect. Committing oneself to a moral profile does not commit a theorist to any views about the relationship between entities with agency and without, the causes of racist outcomes, or the most appropriate or effective solutions. In other words, one is open to explanatory pluralism.

For my purposes in this chapter, the most important kind of pluralism is entity type pluralism because the individualist versus structural debate turns on entity type monism. Because the purpose of this chapter is to argue against the enduring divide, and both sides constrain the kinds of things that are the fundamental bearers of the predicate 'racist', I aim to show the relationship between different kinds of pluralism in order to highlight that neither individualists or

---

[22] A notable exception here are the structuralists responses to Madva's "Anti-Anti Individualism" on the Brain's Blog in which some interventions pluralism began to flourish among structuralists. I see this as a promising way forward, and I hope my levels of pluralism frame adds to this direction.

structuralists are entity type monists. Ultimately, I hope to convince you that we should be *entity pluralists* about the predicate 'racist'.

In this section I've identified four questions we'll want to answer in giving our theories of racism. Further, I've claimed that the questions are nested. In other words, I've claimed that answers to "bigger" questions can obscure our thinking about "smaller" questions if we don't fully appreciate the ways in which answers to "bigger" questions are dependent on smaller questions. As a result, answering "big" questions without identifying the normative core through answering questions of moral status, can unnecessarily and inaccurately constrain our answers to the "small" questions. The primary result of this constraint I'm worried about in this chapter is entity monism. My ultimate goal in this chapter is develop tools for understanding and to motivate entity pluralism. In the next two sections, I'll demonstrate that individualists and structuralists are committed to entity monism and show exactly how things go awry when these questions are _unnested_.

2.4.3   Moral Status and the Predicate 'Racist'

Before moving onto an analysis of individualists and structuralists and my case for entity type pluralism, there are a few basic clarifications that need to be made: What's the relationship between accounts of racism and the concept RACISM? How is the concept RACISM related to the predicate 'racist'? And what role does the predicate 'racist' play in our theories of racism? The most important take-away from Haslanger's discussion is that accounts of racism or theories that answer the question, "What is racism?" are projects primarily concerned with the concept RACISM. In other words, accounts of racism are articulations (whether conceptual, descriptive, or analytical) of the concept RACISM.

This will come up again in this chapter, but for now, it is useful to mention the different projects. Blum (2002, 2014) seems to be engaged in a conceptual project that is at times analytical. When he suggests that we ought to limit our use of racism to its historical or everyday usage, he is engaged in a conceptual project (a project concerned with how people actually use the word in ordinary language), but there are times when his articulation is motivated by the goal of effectively holding others morally responsible, which is a concern of an analytical project. Mallon (2016) is engaged in a descriptive project regarding human-kind cognition (he articulates only one small

part of a theory of racism). And Haslanger (2012a) identifies herself as being involved in an analytical project, as are most other theorists.[23]

We've established that theories of racism are articulations of the concept RACISM. Depending on what theory of concepts to which one is committed, the exact details of what it means to articulate a theory is going to change. But, we can put it roughly: The concept RACISM has some contents which points to some stuff in the world. The stuff in the world that falls under the concept RACISM will have lots of different properties, some of which they will have in common and others of which they will not. The properties they have in common is what allows lots of different stuff to be picked out by the concept RACISM. Part of what articulating the concept RACISM will consist in is identifying the properties that the stuff in the world that the concept RACISM points to have in common.

For the purposes of this chapter, I will assume that all things that are correctly referenced by the natural language term 'racism' have the property of being *racist*, which is identified by applying the predicate 'racist'. Thus, all accounts of racism will have an (in most cases, implicit) account of the application conditions for the predicate racist, i.e. a list of conditions that fix when we should apply the predicate to an entity.[24] Note, however, that application conditions for the predicate 'racist' alone will not constitute a new theory of racism. They won't, for example, say much about whether or not races exist or give us an explanation of why races (real or not) influence our thoughts and judgments. This is perhaps why application conditions for the predicate 'racist' have not been the primary focus in accounts of racism. However, as we will see, this has contributed to accounts of racism failing to account for all the kinds of things we want to capture with our concept of racism.

---

[23] Another feature of analytical accounts that is not clearly identified by Haslanger here is that they often try to explain injustice or oppression with a *unified* causal story. This is because their theories are motivated by justice and are ameliorative projects. In other words, in developing their theories, they hope to provide clear (read: unified) inroads for intervention that will change the world for the better.

[24] Barnes (2019) has argued that metaphysical accounts of gender need not fix the application conditions for gender terms, but that these accounts *should* guide our use of gender terms. Thus, on her account, it isn't the case that all conceptual articulations include the application conditions for terms like predicates. However, she articulates this position against the tradition of assuming application conditions come for free or alongside metaphysical accounts, so my assumption here that application conditions for the predicate 'racist' come along with articulations of the concept of RACISM is not unprecedented.

2.5   Entity Monism

In section 2.2, I reviewed Harris' (1998, 2018) taxonomy for accounts of racism and reviewed his argument for why logical-explanatory accounts of racism fail. My purpose in reviewing this argument is to motivate initial concerns about existing accounts of racism. In sections 2.3 and 2.4, I introduced a new frame for understanding differences between individualist and structuralist accounts of racism. I call this frame *levels of pluralism*. In service of understanding this frame, I've identified four project motivations (descriptive, evaluative, mechanism identification, and ameliorative) with four broad sets of questions for each motivation (moral status, entity type, causal, and interventions focused). In this section, I will come back to the distinction between individualist and structuralist accounts of oppression (including of racism). My goal in this section will not be to try and "fix" the individualist versus structuralist debate. Rather, I'll lay the groundwork for showing how accounts on both side of the debate go wrong. Specifically, I'll argue that these views are committed (implicitly or explicitly) to entity monism.

Entity monist views (with respect to racism) are views that restrict applications of the predicate 'racist' for particular kinds of entities. In other words, the accounts argue or assume that some entities (individuals or structures) bare the predicate 'racist' in some special, fundamental way. Though most theorists do not have an explicit theory for when and how the predicate 'racist' should be applied, they do signal that they think some things are *really* racist while other things are *unimportantly* or *merely consequentially* racist. This is often signaled by phrases like "x kind of thing is of primary moral concern." Blum (2014, 57) also notes this trend writing, "When a plurality of manifestations of racism is explicitly acknowledged…the plurality is seen as either directly derivative from or secondary to a core form of racism." Many theories of racism identify an entity or set of entities in which the primary moral wrongness of racism is located. These views are committed to *entity monism* (I'll argue against entity monism and for entity pluralism in Section 2.6).

2.5.1   Individualist Accounts

In this section, I'll discuss several candidates for a theory of racism that offer individual or personal level explanations (Appiah 1990, Blum 2004, Garcia 1996, 2004, Hirschfeld 1996). On these accounts, what's wrong with racism is in the heads and hearts of individuals. For example, Garcia (2004, 41) writes, "racism's central forms lie in what a person wants for those assigned to

a racial group, and in how she feels about them, what she hopes for them or aims to do to them, and so on." Though the folk understanding of racism often includes references to personal beliefs and behaviors, these individualist views are in the minority within philosophy, but perhaps dominant in society as a whole (empirical work would be needed to know).

But, there is this sub-set of philosophers that haven't been convinced by the structural argument, particularly when it comes to terms like racism. They argue that though there may be structures that lead to inequitable outcomes, this is not where the *wrongs* of things such as racism lie. Rather, these isms are the bad attitudes of individuals. Garcia (1996, 6) takes racism to be a type of moral disregard for people *qua* race.[25] For Garcia, racism is in the heart. Lawrence Blum (2002, 210), on the other hand, argues that racism is in the head; racism is treating and viewing an individual as inferior or harboring dislike and even hostility for an individual because of their race. Appiah (1992, 13-14) argues that racists believe that there are morally relevant differences between members of different races and behave in ways that reflect these beliefs (e.g. by treating members of one race as inferior to members of another). Blum (2002) is quick to note that there are all sorts of racial ills about which individuals should be concerned that occur even when individuals don't have these beliefs, but these racial ills [26] are not racism. [27] Due to their characterization of racism, both of these philosophers are most often characterized as individualists.[28]

In many ways, Garcia's view is much stronger than Blum's. Even though they are both individualists in relevant senses, Garcia argues that the structural is reducible to the individual, where Blum thinks that structural-level concerns are better categorized as a social ill of a different type than racism. Garcia (1996, 11) writes,

> …racism within individual persons is of primary moral and explanatory import, and institutional racism occurs and matters because racist attitudes (desires, aims, hopes, fears, plans) infect the reasoning, decision-making, and action of individuals not only in their private behavior, but also when they make and execute the policies of those institutions in which they operate.

---

[25] He notes that this moral disregard need not be conscious, allowing room for implicit bias (at least on some accounts). See Faucher and Machery (2009) and Garcia (2011) for a debate.

[26] Garcia (2011) also allows for various racial ills that do not count as racism, though he suspects most of the ones we may want to identify have their roots in moral disregard and a lack of care for the welfare of a person qua race.

[27] Blum's view comes from a worry that racism as a concept has been inflated such that it applies to all objects of moral concern having to do with race. He worries specifically that these means we only have two options: label an individual as racist or say there is not problem. He argues there needs to be another way of criticizing individuals with respect to their racial attitudes without labeling them as racist.

[28] See also Hirschfeld (1996).

For Garcia, changing people's hearts is the key to both eliminating racism and various racial inequities. Because of the explanatory and moral power of individual attitudes, the individual is ought to be our primary concern.

2.5.2   Structural Accounts

Garica (2004, 39-40) argues that structural accounts are unable to address or label as immoral beliefs of individuals that do not seem to participate in the larger system of racial oppression; either they are unable to contribute meaningfully to this system or they do not invoke it as an explanation for the belief. He uses the fictional character Shaylee Ledbetter as an example of someone with beliefs that we would like to label as racist. Garcia argues that structural accounts of racism do not give us the framework for doing so. He argues that, rather than identifying the moral wrong as in the individual as ill-will or disdain, structural explanations put the blame outside of individuals. However, it seems that the motivations for doing so are often compelling. There seem to be important cases in which there is no (conscious) racial hatred or disregard but there are negative outcomes along racial lines. In order to test the merits of these views, I'll move to some examples of structural accounts (Cole 1997, Haslanger 2015, Mills 1997, Shelby 2014, Yancy 2008).

Perhaps the most compelling argument against purely individual psychological explanations of racism, are those that demonstrate that injustices along racial lines may occur even when none of the individuals in the scenario are harboring negative racial beliefs. So, in virtue of the way the world has been built and institutions have been structured, injustice occurs. Philosophers who prefer structural explanations over individualistic explanations often charge those working on implicit bias or with volitional/cognitive accounts of racism with making the assumption that a particular instance of injustice would not have occurred if that person had not been implicitly bias or held those beliefs (Alaya and Vasilyeva 2015, Haslanger 2015). Because they think this assumption is mistaken, they argue that the focus on individual accounts, such as those (allegedly) given in accounts of implicit bias and its effects, leads us away from the root causes of social injustices (Banks & Ford 2009, 2011; Haslanger 2015).

The most extreme structuralist approach within philosophy is Elizabeth Anderson's (2010) views on the moral imperative of segregation. According to her view, the root and fundamental cause of racial inequities can be linked to *de facto* segregation. Though there may be other

supporting causes such as limited public transportation and norms around hiring and job advertisement, these issues are structural as well. Similarly, Haslanger (2015) argues that the primary cases of injustice for which we should have concern can be explained without appeal to the mental states of individuals and with only the structural factors. She gives three fictitious cases in which the structures of social life, the schemas under which we operate, and the resources that are allocated according to policies limit individual's options such that unequal or inequitable circumstances obtain. In these examples, no one person did anything wrong; rather, the structural or institutional factors cause and maintain injustice.[29]

2.5.3   Cultural Accounts

For the purposes of moving forward with my entity pluralist account and arguing against entity type monism, I've highlighted the debate between individualists and structuralists. However, there are some limited cultural theories of racism. Though these accounts move us forward in deconstructing the dichotomy posed by the individualists and structuralists (as we do in Davidson and Kelly (2018)), they are still often guilty of entity type monism. For example, Witt (2011) who provides an account of gender (but not of race) provides a cultural account of women's oppression based on social roles and social norms. However, she also claims (like Haslanger 2016) that discussions are often too focused on individuals and individual psychologies. She (2011, 128) writes, "Gender uniessentialism [Witt's view on gender] directs our attention away from individual psychologies, their conscious and unconscious biases, and 'deformed' processes of choice, and toward the social world, its available social roles…" So, though Witt does a nice job of breaking down the dichotomy between individualists and structuralist explanations (Davidson and Kelly 2018), she maintains an entity type monist kind of position; it just happens to be between individuals and structures. Indeed, it may be more appropriate, on the typical division between individualists and structuralists, that she be labeled as a structuralist since these accounts can include both formal and informal structures.

In this section, I've shown that individualists and structuralists are entity type monists. Individualists begin with entity type monism, and structuralists are entity type monists in virtue of their causal monism. In taking this strategy, both kinds of views are subject to Harris' (1998, 2018) critique of "racism as a logic" accounts. One of the primary problems Harris (2018, 40) identifies

---

[29] Garcia (1996) thinks that claims like these include temporal confusions about causes.

with these accounts is that they map the moral wrongness of racism onto the mechanistic explanation provided. For these views, what makes something racist wrong is that it was caused by racism. This direction of fit makes logic-based accounts of racism entity monist accounts. I'll use my account of Nested Questions in the next section to show why this is the case.

2.6   Unnested: Why Individualists and Structuralists are Mistaken

In this section I demonstrate what goes wrong when the four questions I've identified are *unnested*. In other words, I'll show how answering questions in a way that is insensitive to the dependence relations between them leads to mistakes. Specifically, unnesting causes both individualists and structuralists to be entity type monists, claiming that only certain kinds of entities are racist. Individualists go wrong because they let entity type questions overshadow questions of moral status, and structuralists go wrong primarily because they allow causal questions to overshadow questions of moral status. These errors are motivated by praiseworthy goals. Specifically, individualists are evaluation-oriented in that they think answers to entity type questions will allow us to morally evaluate others and indicate that they are responsible for wrongs, and structuralists are mechanism identification- and amelioration-oriented in that they think answers to causal questions will allow us to explain instances of racism in a way that allows us to come up with interventions.

2.6.1   Evaluation-Oriented Individualists

Let's start with the individualists. Garcia (1996, 6, 9) writes, "Racism, then, is something that essentially involves...our wants, intentions, likes, dislikes, and their distance from the moral virtues," and "We can also allow that an action is racist in a derivative and weaker sense when it is connected to racist disregard." Here, Garcia is limiting the kinds of entities that can be fundamentally racist to a particular set of entities, specific affective states of individuals. The first question he answers is a question of entity type: What entities are involved with the wrong? This is before he identifies what's wrong with racism or what features racism's moral profile might share.

Beginning with entity-type questions focused on evaluating individuals and their attitudes as racist or not, causes individualists to focus on individuals (on their beliefs or disregard) and to exclude a whole host of things we might want to describe as racist (norms, procedures, policies,

laws). This focus makes him an individualist. Further, it causes them to be entity type monists. This is caused by answering the "bigger" entity type questions before questions of moral status. Thus, the kind of entity being evaluated constrains the kinds of entities that can come under evaluation as racist, making the individualists individualists, or a kind of entity type monist. I do not think that individualists are totally misguided. In focusing on the entity type questions, individualists are *evaluation-oriented* in that they think answers to entity type questions will allow us to morally evaluate others and indicate that they are responsible for wrongs.

Blum (2002, 2007, 2014) is often taken to be a thoroughgoing individualist. Unlike Garcia, Blum identifies beliefs as the primary entity involved with racism. However, he also claims that racism can be found in many locations and that many different kinds of entities can be racist.[30] He (2002, 208) writes:

> A different source of confusion and moral overload regarding racism concerns what one might call racism's location. Many different kinds of entity[sic] can be racist – actions, institutions, practices, symbols, statements, jokes, persons, to name a few.

So, though Blum locates the primary moral wrongs in the affects and beliefs of individuals (seemingly constraining by entity type), it may be better to understand Blum as answering causal questions first (which is more common with structuralists). He (2002, 211, italics added) suggests that "we confine racism to *manifestations* or *representations* of racial antipathy or racial inferiorizing." Thus, on Blum's (2002) account, entities (like jokes) are racist because they are *manifestations* of racial hatred or beliefs about racial superiority. In other words, that a joke is racist is caused or explained by racial hatred or racist beliefs. On either interpretation of Blum's view—as answering entity type questions first or causal questions first—his answering of "big" questions first leads him to leave some things out of his analysis. And, this shouldn't be surprising, since this is in fact his goal (with the ultimate goal of avoiding conceptual inflation and preserving the moral function of the predicate, but, as will be made more clear in Chapters 3 and 4, this can been achieved without drastically constraining the concept of racism).

By allowing their answers to entity type questions to pervade their theories, individualists exclude some things that should be included in our moral theories of racism. In addition, notice that constraining moral status by entity type alone is very difficult. For example, imagine an individual tells a racially-disregarding joke that uses both a racial slur and relies on inaccurate

---

[30] One possibility here is that Blum is either inconsistent or his views have developed over time.

stereotypes within the context of a group that has norms against intervening or calling out these kinds of jokes that is currently located inside of an institution that has no formal policies or rules against telling these kinds of jokes. If I'm an individualist like Garcia (1996), who answers questions of entity type first, I'll assert that the individual who told the joke is racist (or, if we want to get more fine-grained as my account suggests, his racial disregard is racist). Or, if I'm Blum (2002), I'll say that the joke was racist only if it is a manifestation of an individual's racial hatred or superiority. However, just as Haslanger (2016) argued that focusing the implicit biases of individuals doesn't tell the whole story, these analyses leave out the norms of the group and the formal policies of institutions. If, however, we shift our analysis to be first focused on the causal features, we may lose out on ways to hold the individual accountable. After all, perhaps this individual has the power to institute a formal policy against racist jokes and holding him accountable would prompt him to do so (see Madva 2016 for why individuals matter to structural change).

2.6.2   Mechanism Identification-Oriented Structuralists

Structuralists suffer a similar fate, but for different reasons. Haslanger (2016, 2) writes, "The normative core of what is wrong with racism/sexism lies not in the 'bad attitudes' of individuals but in the asymmetrical burdens and benefits and inegalitarian relationships that societies impose on such groups." Though Haslanger doesn't say that individual attitudes cannot be racist, she argues that there is something special about structures that cannot be captured by talking about individuals. This is because she thinks what best explains the outcomes we're concerned with in giving our accounts of racism (or sexism) are structures. On Mills' (2011) racial contract view, the racial contract explains racism. Thus, if there is a wrong that isn't explained by the racial contract, then it isn't racism. On Shelby's (2014, 66) racism as an ideology view, "Racism is a set of misleading beliefs and implicit attitudes about "races" or race relations whose wide currency serves a hegemonic social function." Shelby's (2003, 158) account counts as a structural account because on his view ideology is not primarily about beliefs held by individuals but that the beliefs are widely shared in a given population and that others in the population know that they are widely shared. Similarly to Mills, all racist things must be explained by the ideology and the function it plays in order to be counted as racist.

Each of these views is motivated by identifying a mechanism that causes pernicious outcomes. However, rather than first describing these outcomes and identifying what is wrong with them, these theorists begin with this causal story. Beginning with causal questions focused on explaining why there is racism and what the interventions for interrupting and ending racism should be, means that structuralists focus on structures (policies, procedures, laws) and exclude a whole host of things we might want to describe as racist (beliefs, ill-will, individual actions, norms). This focus makes them structuralists. Further, it causes them to be entity type monists. This is caused by allowing answers to the "bigger" causal questions to constrain their answers to the "small" questions of moral status *or* of entity type.

The causal story they tell about racism (whether it be the asymmetrical burdens imposed by society, the racial contract, or ideology) constrains the kinds of entities that can come under evaluation as racist. Their causal monist makes them entity type monists as well. This is the defining feature of structuralists. Again, I do not think that structuralists are totally misguided. In focusing on the causal questions, they are mechanism identification- and amelioration-oriented in that they think answers to causal questions will allow us to explain instances of racism in a way that allows us to come up with interventions. This is certainly a good motivation, but it leads to both metaphysical and descriptive problems.

The under appreciation of the dependence relationship present within the nested questions causes individualists and structuralists to exclude some entities from being evaluated as racist that should be. Both families of views are tacitly committed to entity type monism and this comes out because they unnest the questions that must be answered to give an account of racism. So, what happens when we _renest_ the questions? When we appreciate that moral status should be at the core of our accounts of racism, determined before a causal mechanism or a plan for intervention, we are able to focus on a wide variety of entities (unconscious attitudes, sub-personal states, avowed beliefs, norms, cultural practices, policies, laws, and institutions) without excluding things we might want to describe as racist. The result is that one is an entity type pluralist, not an entity type monist, neither an individualist nor a structuralist.

2.7   Conclusion

In this chapter I have provided a new framework, levels of pluralism, for understanding accounts of racism. I've developed these tools in order to diagnose the enduring debate between

individualists and structuralists, to show that both of these views include a shared mistaken metaphysical commitment. One thing we learn with the introduction of these tools is that individualists and structuralists have more in common than we may have previously thought, and consequently, a whole host of other views (explanatory pluralist views) have been left out of the philosophical picture. Without these tools for identifying locations of agreement, disagreement, and new possibilities, the debate endures because we are unable to adequately argue against either side. With these new tools in hand, the debate is further deflated, and a new pathway can be forged.

One upshot of my commitment to entity type pluralism, and the view that descriptive questions of moral status should be the core or inner-most feature of our accounts of racism, is that many entities can be evaluated as racist (from implicit attitudes, to police violence, to mass genocide). This is something that individualists and structuralists desperately try to avoid for a whole host of reasons depending on their priorities and motivations. One reason that I haven't mentioned is worries that we will take all instances of racism equally seriously. For example, Blum (2014, 58, emphasis added) cautions against taking outcomes that are of "relatively less moral weight than core forms of (what is generally understood as) racism," and "inflat[ing] their *moral significance* by implying that the opprobrium generally attached to unquestioned instances of 'racism' applies *equally* to less racial ills." Of course, it matters here what Blum means by both "moral significance" and "equally," but we can understand the basic thrust of his worry.

To avoid this, Blum (2014) claims that we ought to attempt to explain moral wrongs having to do with race using terms other than 'racism' first. However, if one of the primary goals of conceptually analyzing racism is to figure out what makes racism wrong or what moral features racist outcomes might have in common, then it seems question begging, or at the very least bizarre, to sort the morally problematic outcomes having to do with race into "really bad stuff" and "sorta bad stuff" and to consequently *exclude* from the analysis the "sorta bad stuff" before one begins. But, this is exactly what Blum insists on. Blum (2014, 63) seems to think that labeling two things *racist* implies "a kind of moral equivalence." However, just like things can be *wet* or *very wet*, the property *racist* comes in degrees. Thus, though my view certainly goes in the direction that is worrisome to Blum, it neatly avoids the conceptual issues of "pre-sorting" into moral significance categories before even starting the moral analysis.

However, there's one thing I'd like to bring forward from Blum's analysis. As much as I disagree with Blum's project of limiting the concept RACISM and the use of the predicate 'racist'

to only the most pernicious, historically-consistent aspects, I wholeheartedly agree with a conclusion that comes from his analysis of cases. He writes (2002, 216) that we ought to "encourage clarity as to the location of what is, or was, racially objectionable in the situation." In other words, when evaluating a situation, we should be more fine-grained in our analysis. Rather than evaluating a person as racist, we should evaluate their action (or their mental state, cognition, assumptions, participation in an activity, etc.) as racist. The introduction of this clarity around the location of racism will help us with accountability and intervention. In the next chapter, I provide an account of the application conditions for the predicate 'racist' that takes up this strategy. Then, in Chapter 4, I provide an account of intervention-sensitive moral responsibility that allows agents to be held responsible for intervening on non-agential racist entities. And, finally, in Chapter 5, I discuss interventions into our epistemic resources that will help us see more clearly when the application conditions for the predicate 'racist' has been met and consequently when we have the responsibility to intervene. It is my hope that these chapters work together to increase clarity around the concept of RACISM in a way that preserves the moral function of the predicate 'racist' so we can hold each other accountable.

CHAPTER 3.   ENTITY TYPE PLURALISM IN ACTION: APPLICATION CONDITIONS FOR THE PREDICATE 'RACIST'

3.1   Why Application Conditions?

People often use the predicate 'racist' to claim that something bad involving race has happened. In an attempt to describe these wrongs, they may exclaim, "That's racist!" In this way, individuals adeptly use the predicate 'racist' in a wide variety of situations with vastly different features. For example, an individual might use the predicate to identify what's wrong with the use of a racial slur or a police officer shooting an unarmed Black or Brown individual or to explain why testimonies of Women of Color go unrecognized or unheard, why the pool of applicants for an academic position in philosophy are all white, why People of Color are inequitably paid, or why schools in communities of Color meet fewer educational outcomes. Though these cases vary along several dimensions, the use of the same predicate signifies that we think the cases have at least one property in common. However, there are many instances of disagreement, too. Someone may claim that something is racist, while others dissent, claiming that whatever has happened is not actually racist, but only allegedly or seemingly so. These disagreements have more than merely academic importance because 'racist' is used to express a significant moral appraisal and as a tool for accountability. In other words, when we use the predicate 'racist', we often do so not only to describe but to change people's behaviors.

In the previous chapter, I argued for entity type pluralism or the view that our theories of racism and our application of the predicate 'racist' shouldn't be reserved for particular kinds of entities. I argued for this position against the background debate between individualists and structuralists. I identified both broad families of views—individualists and structuralists—as entity type monist views that argue certain kinds of entities are the fundamental bearers of the predicate 'racist'. Now that the ground for entity type pluralism, a view that is neither individualist or structuralist, has been cleared, I've opened space to articulate the application conditions for the predicate 'racist' in a way that is not constrained by entity type. Though the entity type pluralist view helps us understand that we shouldn't reserve the application conditions for certain entities, it by itself doesn't specify when and why we should apply the predicate to different entities. The goal in this chapter is to provide such an account.

To do so, I delineate and argue for a specific set of application conditions for the predicate 'racist'. These conditions tell us when it is correct to apply the predicate to a particular entity in a particular case. Ideally, the application conditions for the predicate perfectly capture the extension of the predicate, and so an account of the proper application conditions will allow us to adjudicate debates about what is genuinely racist. It will also, however, preserve and shed light on the moral functions that the use of the predicate allows us to perform. [31] The pluralist account of the application conditions for the predicate 'racist' that I develop satisfies these desiderata, but it also challenges major conceptions of racism in the literature. The application conditions that I develop in light of entity type pluralism will show that there are many kinds of entities (e.g. unconscious attitudes, sub-personal states, avowed beliefs, norms, objects in the world, cultural practices, policies, laws, and institutions) to which the predicate 'racist' can be applied and, moreover, that each type of entity qualifies as a fundamental bearer of the property picked out, such that its status of being racist need not, and in many cases cannot, be reduced to or explained by appeal to its connection to some other, more fundamentally racist, entity. On the entity type pluralist view, there is no one fundamentally racist type of thing. There is no single dark heart of racism.

In Chapter 2, I talked a lot about accounts of racism, what giving a full account requires, and developed tools for distinguishing between and analyzing different accounts. In developing these tools, I argued that our accounts must be sensitive to the dependence relationships between sets of *nested focus questions*. I showed how answers to "bigger" focus questions can obscure our thinking about "smaller" questions when this dependence relationship is not taken seriously. In particular, I think there are good reasons to believe that answers to the other kinds of questions— questions seeking to offer a *unified explanation* for racism— obscure thinking about the application conditions for the predicate. Thus, it should not be surprising that I'm starting with the first set of nested questions, questions of moral status, in offering the application conditions. As I

---

[31] An assumption here is that an accurate semantic account of the predicate 'racist' will help guide a pragmatic account of the effective uses of the predicate 'racist' within our practices of holding each other accountable. In other words, there's an assumption that getting a good metaphysical account on the table might give us insight into *what is being said* when we communicate moral appraisals to one another. In contrast to the assumption I take on in this chapter, Blum (2002) argues for the opposite direction of influence. For him, pragmatic concerns about communicative practices guide his metaphysical account of what racism is. This feature makes his view a mixture of what Haslanger (2012) calls *conceptual* and *analytical* projects (refer back to Chapter 2 for more details). A few more details of this methodological commitment will be laid out in Section II of this Chapter, and my account of accountability that flows from this semantic account will be given in Chapter 4.

suggested in the previous chapter, the predicate 'racist' tracks the morally relevant property within accounts of racism.[32]

Here's a concrete example about how questions of explanation can obscure our thinking about the predicate 'racist': when the average person hears that races don't exist, i.e. that there are no biological essences that track our contemporary use of racial categories, they may conclude that there can't be anything that is racist. Given our dominant conceptual resources, this isn't a completely irrational conclusion. However, it is mistaken. By focusing first on the application conditions for the predicate, I hope to avoid these kinds of mistakes, as well as the others detailed in Chapter 2. But, in giving a set of application conditions, I will not offer a new theory of racism. A full theory of racism will need to include some sort of account of the application conditions for the predicate,[33] but it will also require many more claims. In the previous chapter, I gave a stronger and more detailed argument for why it is useful to conceptualize questions of moral status as the core, the inner-most feature, of accounts of racism. So, I delineate the application conditions for the predicate 'racist' separately from the "bigger" questions of entity type, causal, and intervention-focused questions.

Here's how I've organized this chapter: In Section 3.2, I'll give a brief defense of the method for this chapter of the project, which I'm calling *epistemic position engineering*, and separate this kind of project from other kinds of projects, e.g. mere conceptual analysis or projects that attempt to identify how a term is *actually used* by the average person or *should be used* given certain aims.[34] Improving our epistemic position with respect to the concept of RACISM will tell us how and why we should apply the predicate to different entities.

---

[32] It is important to note here that I approach this project not as a philosopher of language, but as philosopher of race, mind, and social epistemology. I've chosen to articulate my project this way because of the philosophical and popular debate about when and why we should call something racist. I've found that thinking about this question through the lens of concepts, meaning, and reference is fruitful for clarifying disagreements and providing a solid foundation for the kind of claims we want to make with respect to the predicate 'racist'. An equally important motivation is that in the popular conversation around 'racist', many of the folks who want to limit the use of the predicate are white. I view this as a strategy for preserving "white goodness" and avoiding accountability for racism.

[33] I say some sort of application conditions here because most (if not all) existing theories do not explicitly discuss application conditions for the predicate 'racist'. Rather, the application conditions are assumed to fall out of a larger theory of racism. The account I provide, thus, increases clarity about the kinds of things and the situations in which we can appropriate apply the predicate.

[34] For example, as discussed in the previous chapter, Haslanger (2012a) gives what Mallon (2016, 209) identifies as a "justice-driven metaphysics" for the social categories of race and gender. Haslanger's account of these human kinds is developed with certain aims in mind, e.g. eliminating or mitigating systematic subordination. Though my account of the predicate 'racist' shares some features of this strategy in that I think delineating the application conditions for the predicate will be helpful for our moral aims, my strategy is not to employ moral evaluative aims in order to

In Section 3.3, I start with Harris' (2018) actuarial account of racism and build from there to identify the application conditions for the predicate 'racist'. Utilizing empirical data on life expectancy and health outcomes, I'll give an argument that these application conditions map onto the extension of the predicate 'racist', which expresses the concept RACIST, which in turn locks onto the property *racist* in the world.[35]

In Section 3.4, I use the application conditions to explore a case study and identify some potential problems for the account, specifically with respect to disagreement, that will help us understand the kinds of questions a framework for accountability might need to answer. Because my account will show that we can appropriately apply the predicate 'racist' to non-agential entities, we'll need some way of identifying the relationship between agents and these entities. I'll give a full account of accountability in Chapter 4.

My goal in this chapter is to clarify the concept of RACISM, especially the predicate 'racist', to provide a pathway for improving our use of natural language in a way that allows the predicate 'racist' to more effectively express moral appraisals and to better contribute to our collective practices of holding each other accountable.

3.2   Epistemic Position Engineering

In Chapter 2 I offered a new frame, *levels of pluralism*, for understanding what different accounts of racism are up to. In this chapter, I offer a new methodological framework, *epistemic position engineering*, for understanding the project of delineating application conditions. Through this method, we can improve our epistemic position with respect to a particular concept. I articulate this methodology first because this project departs from the three kinds of projects articulated by Haslanger (2012a) and discussed in Chapter 2. First, I'm working from a broadly externalist metaphysics of meaning and theory of concepts. Though the formulation may shift slightly depending on the metaphysical framework one might choose, none of my arguments depend on the complex distinctions and disagreements about reference relations between natural kind terms

---

delineate the application conditions. Rather, I think delineating the application conditions independently of our moral aims will ultimately assist us with our moral endeavors.

[35] Note another difference here between myself and a philosopher of language: I want to give a semantic account of the predicate 'racist' that is independent from a pragmatic account of what's going on when the word is used to hold someone accountable.

and concepts.[36] This commitment constrains and informs what I think is happening when we identify application conditions for the predicate 'racist': we improve our epistemic positions with respect to the concept RACIST.

In delineating the application conditions of the predicate 'racist', I hope to improve our epistemic position with respect to the concept RACIST, so we can identify the property of *racist* in the world in order to avoid misapplications and mistaken non-applications of the predicate 'racist'. I will call this improvement project *epistemic position engineering*. This project is distinct from conceptual engineering in its metaphysical commitments. In contrast, within a framework of conceptual engineering, we assume that our use of predicates determines the content of our concepts. *Epistemic position engineering*, on the other hand, is a project of improving our *epistemic position* with respect to a concept.[37] Delineating the application conditions for the predicate improves our epistemic position by giving us insight into the semantics of the predicate. Once we have improved our epistemic position, we might find that we need to change our use of the predicate so that it better reflects the semantics.

In improving our epistemic positions, we are better able to express the concepts we think we are expressing in our use of natural language to pick out the properties in the world that we are interested in identifying. Through a project of *epistemic position engineering*, we can come to know more about a property we care about, in this case *racist*, by thinking carefully about the application conditions for the predicate. Though the framework of this project does assume that there is a property "out there" to be picked out by a particular predicate, it does not assume that these properties exist necessarily or ahistorically, but rather contingently.[38] Thus, the application conditions that will be identified here are situated within a complex historical picture laden with contingent realities.

The method of *epistemic position engineering* is also different from determining how people in the actual world apply the predicate 'racist'. Though I am interested in and will often refer to the ways in which we currently use the predicate 'racist' to capture and morally appraise

---

[36] I endeavor to consistently formulate my arguments such that they are consistent with an externalist view of content fixation, and I further attempt to be consistent with Fodor's non-inferential concept individuation, see Fodor (1990, 2004) for more insight.

[37] This project is fraught with metasemantic issues. I take myself to be operating within a roughly causal theory of reference with some revisions (see Haslanger 2012b; Burgess and Plunkett 2013a, 2013b; Plunkett and Sundell 2013; Mallon 2017; Schroeter and Schroeter 2018 for more insight into the debates within metasemantics).

[38] Modal contingency for social kinds is consistent with a causal-historical theory of reference for social kind terms.

specific phenomenon as a starting point, I ultimately want to offer a set of application conditions for the predicate 'racist' that will likely revise some of our current practices of applying the predicate. There are several reasons that motivate the revisions I suggest and some of the motivations are not purely epistemological, i.e. the project here is not only about getting our representations of the world right, but to get the world right in order to use language in a way that assists us in improving the world. An assumption in this project is that getting the world right (improving our epistemic positions), will also improve our moral positions (our ability to be morally evaluated and held accountable in the right kinds of ways such that there are fewer racist things in the world).

3.3 Application Conditions for the Predicate 'Racist'

In this section I will identify the application conditions for the predicate 'racist' and argue that these application conditions map onto the extension of the predicate 'racist', which expresses the concept RACIST, which in turn locks onto the property *racist* in the world. As was discussed at length in the previous chapter, philosophical accounts of racism can be divided into two *broad families* of views: individualists and structuralists. These accounts typically argue that one set of things—either individuals or structures—are the primary source of the wrongness of racism or the cause of racist outcomes and that all other sources or causes are derived from or reducible to the primary source. Many of these accounts share, either explicitly or otherwise, a commitment to *entity type monism* or the exclusivist clause: *only* entities of this sort can be the primary bearers of *racist*. In this chapter, I'll assume the arguments from the Chapter 2, rejecting entity type monist accounts that identify only one set or kind of entities as possibly racist as reductive because these accounts fail to capture all types of things we'd like to count as racist for both descriptive (i.e. they get the world wrong) and normative (i.e. they keep us from improving the world) reasons.

My argument for entity type pluralism clears the ground for starting with questions of moral status. Questions of moral status focus on happenings in the world and whether the happening was bad. With respect to racism, we use the predicate 'racist' to express the concept RACIST, which shares exemplifications with the morally relevant property in the world: *racist*. So, in order to move us forward in answering questions of moral status, a set of application conditions for the predicate will be very useful. Because I've cleared the ground for entity type pluralism, my account of the application conditions for the predicate 'racist' will show that there

are many entities to which the predicate 'racist' can be applied and, moreover, that each type of entity qualifies as a fundamental bearer of the property, such that its status as being racist need not, and in many cases cannot, be reduced to or explained by appeal to its connection to some other more fundamentally racist entity. On this account, there is no single "primary" source of racism such that other entities that are also *racist* are only derivatively *racist* in virtue of a relation to the 'real' or 'ultimate' source or fundamental bearer of the predicate. So far, I've said a lot of highly technical abstract stuff about predicates, concepts, properties, and application conditions. Some examples will bring us back down to earth (and to the racist stuff in the world).

3.3.1   Is that racist?

I'm interested in the very many kinds of entities that can be racist: unconscious attitudes, sub-personal states, avowed beliefs, norms, cultural practices, objects in the world, policies, laws, and institutions. However, I'll start my analysis with objects in the world, an entity that is often left out of accounts of racism. In his now famous analysis of the bridges over the Long Island parkways, Winner (1980, 121) claimed, "there is no idea more provocative than the notion that technical things have political qualities." In this article, Winner claims[39] that Robert Moses, a New York public official who had lots of power over city planning from 1924-1975, intentionally designed very low overpass bridges over the only parkways leading to Jones Beach in order to keep poor and working-class Black and Puerto Rican people from taking the bus out to the peninsula. Since Moses unapologetically held negative, essentialist views, his motivation isn't hard to imagine. On standard views of racism, it is fairly uncontroversial to call Moses himself racist.[40] Depending on the account, we might straightforwardly call Moses' actions racist, as well as the structure of decision-making that allowed one person with racist motivations to control the bridge height. But are the *bridges* racist?

Winner (1980) says that artifacts can have political qualities, but interestingly the two ways in which things can have political qualities that he discusses are both derivative. The first way is

---

[39] Or, more accurately, reports. The story of the tricky bridge design appears in Caro's (1975) biography of Moses. Though the story is somewhat contested, even those who want to shed questionable light on the story agree that the bridges over the parkway are in fact lower than others built at the same time. One supposed "problem" for the story of Moses is that, at the time, buses weren't permitted on the parkway anyway, bridge clearance problems or not. For an analysis, see: https://www.citylab.com/transportation/2017/07/how-low-did-he-go/533019/

[40] It seems that some structuralists want to argue that Moses isn't racist since he's a person and persons aren't the appropriate unit for evaluation. But, I think in the privacy of these philosophers' homes, they would call it like they see it: Moses is (a) racist.

that the physical ways we arrange things lead to or enforce certain *patterns* of engaging connected to power. In this case, the patterns would be racist. The second is that the artefact could be inextricably *linked* to a system of power. In this case, the system of power would be racist. On Winner's own lights, it doesn't seem like the bridge is racist. These ways don't seem to identify ways that *artefacts* can have political qualities themselves, but rather how they can have political influence in the world. In other words, it seems like the artefacts don't have the property of being political, at least not fundamentally.

However, in this chapter, I argue that the bridges, and other objects, can be racist in themselves rather than in their connection to other things. Our intuitions that *things* cannot have moral properties run deep. This is perhaps because questions and objections like: What if everyone disappeared? Is the bridge still racist? Patterns and systems aren't racist after people disappear because patterns and systems disappear *with* the people. But, the bridge remains with or without people. How is it possible that the properties of the bridge change because something else (namely, the population of all human beings) comes in or out of existence? My answer is that it is not that the bridge's having the property of *racist* is constituted by the people affected by the bridge. It is that the property of *racist* itself only exists if there are people in the world. Thus, we can give an account for how and why the bridge can be racist by itself even if the bridge would no longer have the property without people.

Cases like the Long Island Parkway bridges motivate my account and haven't been discussed much in the philosophical literature on accounts of racism. The special attention given to objects, and their ability to be a fundamental bearer of the predicate 'racist', sets this project apart from the individualist and structuralist accounts that precede it. Though structuralists make a lot more room for closer-to-object stuff, like social structures, they are still *very* human. So, what makes something racist on my account? When do we apply the predicate?

3.3.2 Moral Monism and Harris' Account of Necro-Being

In the previous chapter, I used my *levels of pluralism* frame to argue for entity type pluralism. This allows us to think about all kinds of entities when we give accounts of racism and paves the road to being able to give application conditions for the predicate 'racist' that aren't constricted first by entity type. In addition, I argued that the levels of pluralism (moral pluralism, entity type pluralism, causal pluralism, and interventions pluralism) are in a dependence

relationship with one another, especially as it relates to explanation, which is made up by the latter three forms of pluralism/monism. Commitments to monism within one of these three explanatory levels leads to monism on other levels. For example, as discussed in Chapter 2, the causal monism of structuralists leads to entity type monism.

The moral status level of this framework makes up the descriptive features within accounts of racism. Answers to questions of moral status describe and identify the shared moral profile of the kinds of things in the world we want to capture with our accounts of racism. Because this level captures only descriptive features, rather than explanatory features, the decision to be a monist or pluralist on this level does not influence the monism or pluralism on other levels. Thus, entity type pluralism does not require strong moral pluralism and can accommodate a weak moral monism.

What I mean by this is that the descriptive aspect of an account of racism can identify a particular moral profile, a family resemblance, without committing to any features of explanatory monism (entity type monist, causal monism, or interventions monism). Weak moral monism is not uncommon with respect to racism. Some examples of views that argue for a shared moral profile for racism are racism as domination (Desmond and Emirbayer 2009), racism as hatred (Garcia 1996) racism as disrespect (Glasgow 2009), racism as ideology (Shelby 2003), and racism as upholding a contract (Mills 1997). The application conditions I put forth here will also assume weak moral monism about racism. I'll argue that things that are racist share a moral profile, share in their moral wrongness. However, rather than any of the above features, I'll start with Harris' (2018) Necro-Being account that identifies the moral profile of racism as misery, living death, and morbidity.

As discussed in Chapter 2, Harris (2018) argues that all logical systems of racism that attempt to provide a coherent and consistent set of correlative wrongs for each instance of racism give rise to cases that cannot be explained.[41] Rather than providing a logical (explanatory) account, Harris (2018) articulates a descriptive account of racism in which death is the main indicator of its presence. Harris' account of racism is an actuarial account of racism, or an account that focuses on the provision of death. Like a business actuary who analyzes risks for companies based on

---

[41] In this way, Harris' (2018) critique is similar to Martela's (2017) extension of Dewey's (1930) critique of analytical or logical moral systems that "all postulate one single principle as an explanation of moral life" (quoted in Martela, 61). These theories, which abstract away from the realities of lived experience, attempt to justify foundational moral principles through a priori reasoning alone. Just as the search for a single moral principle ignores the complex realities of our moral landscape, the search for a unified explanatory and logical principle for racism doesn't reflect the racial realities of this world.

particular features (including an ability to predict life expectancy), Harris' account focuses on the risks, the tragic realities, of being negatively racialized in our current world. Necro-being (2018, 39), then, is "a condition that kills and prevents persons from being born." Harris (2018, 57) writes, "The probability of death defines racism: who dies, who benefits from their death, who suffer undue short lives, and who are the targets of life shorting acts." Though this account differs from Mbembé's (2003) bio-political, Foucauldian account of racism as the regulation of the distribution of death, his account shares many similarities.

Harris (2018, 42) uses the concept of tragedy as "irredeemable and meaningless infliction." He uses the concept of tragedy to remain agonistic about (or to reject outright) both metaethical and normative ethical theories. However, as his argument develops, he (2018, 58) writes, "I will take it as given that a moral wrong of racism is the unnecessary race-based sustaining of shortened lives, physical pain, diseases, and premature deaths…" Notice that on this account, race-based death does not have to share a common *cause* to count as racism. Deaths can come about in a variety of ways; hence, Harris' articulation of racism as "a polymorphous agent of death" (1999, 2018). Though Harris' view is clearly explanatorily pluralist, it is morally monist. In identifying death (and various instantiations of death) as the primary wrong of racism, he has articulated a position that does not count other things (often counted as racism on other accounts) as racism. He (2018, 46) writes, "If a racist ill-will has the result of benefiting an individual for example, given my admitted priority of death and health, it does not matter. I will not consider it 'racism'." Thus, though Harris' account is thoroughgoingly pluralist with respect to explanation, it is a weakly morally monist.[42]

In the next several subsections of the paper, I'll bring forth Harris' (2018) account into my application conditions for the predicate racist. My primary motivation for using Harris' (2018) actuarial account of racism is to preserve explanatory pluralism (the combination of entity type, causal, and interventions pluralism). As described by Harris, the primary benefit of *describing* racism rather than *explaining* it is that this method avoids anomalies or things we want to count as racism, but don't fit under our neat explanation connected to the wrongs of racism in a one-to-one correlation. This strategy avoids mistakes by appreciating the dependence relationship within the

---

[42] Another virtue of Harris' (2018) account is that it is from the outset admittedly incomplete. He writes, "This account is not comprehensive, but hopefully its philosophic contours are sufficiently suggestive to recommend viable future avenues of empirical research and sufficiently rich to support the thesis of racism as necro-being" (40). In not claiming to give a full or unified account, Harris leaves open the possibility for moral pluralism.

nested focus questions; we can avoid mistakes by being entity type pluralists or by not reserving the predicate 'racist' for a particular set of entities deemed important for our explanatory framework for how things came to be this way in the world.

3.3.3   Miseries and Premature Death

In this section, I review some of the empirical data on race, health, and life expectancy to show how premature death can be productively operationalized and understood as the descriptive function of racism.[43] I'll start my discussion with the life expectancy at birth for various racialized groups, using data collected by the Center for Disease Control and Prevention (2016) and presented by the Social Science Research Council through their Measure of America Program (2018),[44] which has developed user-friendly tools for exploring data points related to well-being in the United States.[45] Of course, methodological difficulties immediately arise with data related to race. This is especially true with this data, since it is unclear how race is tracked in CDC data collection.[46] Second, the data focuses only on the United States, which is a major limitation. As of this writing, no global life expectancy data coded by race is available.

The life expectancy at birth for African-American people is 74.6, for Native American people it is 76.9, for white people it is 78.9, for Latino people it is 82.8, and for Asian-American people it is 86.5. There are countless factors that go into life expectancy (and I'm not an expert in life expectancy analysis), but I'll say a little bit about how these numbers relate to my thesis. Something to notice immediately is that both African-American people and Native American people have shorter life expectancies than white people. And, these differences cannot be reduced merely to location. The highest life expectancy rate for white people is in the District of Columbia (84.3), which is also the lowest for African-Americans (71.6). In South Dakota, the life expectancy of Native Americans is just 68.2, where it is 80.4 for white people. This straightforwardly supports the idea that racism has something to do with premature death.

---

[43] I say this because many initial reactions to this account is that death is "too high of a bar" with respect to racism. In this section and the next section I will assuage these worries by showing that premature death is often *cumulative* premature death.

[44] See http://measureofamerica.org/maps/ to explore.

[45] I also used a table made from this data. The table can be found at https://www.kff.org/other/state-indicator/life-expectancy-by-re/.

[46] As such, throughout this section I will utilize the categories that have been utilized within the studies without much analysis.

However, both Latinos and Asian-Americans have higher life expectancies than white people. Some (see Scommegna 2013 for an overview) have theorized that the relatively high life expectancy of Hispanic people can be almost solely attributed to the low prevalence of smoking, despite high percentages of the population working difficult jobs with lower incomes. The high life expectancy of Asian-Americans is mostly attributed to lower rates of heart disease and cancer (than white people living in the U.S) (Acciai, Noah, and Firebaugh 2015), which could be caused by a variety of mechanisms. I do not mention these factors to try and explain away differences in life expectancy, but to highlight that life expectancy is indeed complex. Differences in life expectancy among BIPOC[47] may also be attributed to the various ways and degrees racialized groups experience racism. A straightforward way of understanding these differences is that anti-Black and anti-indigenous racisms are more pervasive and life-shortening.

Life expectancy is a good place to start for a descriptive account of racism, but it isn't without its conceptual difficulties. Remember, we're interested in premature death, or death that occurs earlier than it otherwise would've given the absence of racism. Thus, that Asian-American life expectancy is very high compared to other racial groups doesn't mean they aren't affected by racism. Perhaps Asian-American life expectancy would be even higher given the absence of racism. Further, racism is not the only form of oppression that may influence life expectancy. Class, gender, (dis)ability status, sexual orientation may also influence life expectancy in a variety of ways.[48] And, finally, oppression is not the only thing that influences life expectancy—though oppression does constrain the agency of individuals, life expectancy can also be affected by differences in personal choices (even if these choices are constrained by norms, see Chapter 5 for more details).

For the purposes of this chapter, life expectancy data is ideal because it describes a feature of our world, but it doesn't tell the whole story with respect to the descriptive features of health outcomes for oppressed racialized groups. So, I'll now turn to some of the data on morbidity and disease. Black people have higher rates of death due to common diseases. This is caused in some cases by a higher rate of disease (e.g. congestive heart failure) and other times from a higher rate

---

[47] I use BIPOC to refer to Black, Indigenous, and People of Color. This acronym was created to highlight the ways in which Black and Indigenous people experience racism and/or white supremacy in a different way than, for example, Asian people. Some people use this term as to refer to Black and Indigenous People of Color, as a *contrast* to or subset of POC or People of Color. I will not use it in this latter way, though I want to acknowledge the differing patterns of misery, power, and solidarity that this distinction tracks by using the acronym.

[48] As discussed in the Introduction to the dissertation, this is a limitation of focusing solely on racism rather than utilizing a robust intersectional approach.

of death from disease (e.g. cancer). Heart disease is the leading cause of death for all people in the United States; however, the rate of hypertensive heart disease is higher for African Americans, with hypertension accounting for 50% of the difference in morbidity between white people and Black people (Musemwa and Gadegbeku 2017). In 2016, Black people were 14% more likely to die from cancer than white people (Simon 2019). A study analyzing stroke hospitalizations in Maryland found that Black patients were more likely to die or have an undesirable discharge disposition after a stroke (Onukwugha and Mullins 2007).

In addition to higher rates of death, BIPOC often do not receive the same quality of healthcare, even when they have access to it. For example, studies show that BIPOC often have a hard time accessing pain management. Though there is an overall increased in the prescription of opioids for pain-related emergency room visits, white people are more likely to receive opioids than all other racialized groups (Pletcher et al. 2008). Black children are less likely to receive appropriate analgesic treatment (painkillers) when they visit the emergency room, even when the hospital serves mostly Black clients (Rasooly et al. 2014). Black women with endometriosis are more likely than white women with endometriosis to experience delayed diagnoses for the condition (Bougie, Healey, and Singh 2019). Both Black and Latino patients are less likely to receive care that requires a doctor's order than white patients even when they have the same illness or disease (Ashton et al. 2003).

Racial disparities also arise at the beginning of life in the United States. In 2017, preterm birth rates were higher among Black and Hispanic mothers than among white mothers, with rates for white mothers at 9.05%, for Black at 13.93%, and for Hispanic at 9.62%. Low birth weight is one of the leading causes of infant mortality (Heron 2017). Very low birth weights for Black infants (2.39%) were more than three times as high than for white infants (0.77%) between 2006 and 2016 (Womack, Rossen, and Martin 2018). Total infant mortality also varies greatly by race, with Black people having the highest rates—11.11 per 1000 births—compared to 7.61 for Native American people, 5.06 for white people, 5 for Hispanic people, and 4.07 for Asian people. These differences in pregnancy complications and infant mortality contribute to the lower life expectancy and to the prevention of being born.

The United States also has a history of forced sterilization. Compulsory sterilization in the United States began within eugenics programs aimed at ending things like "hereditary insanity" and primarily targeted institutionalized individuals. From the origin of the practice, forced

sterilizations were considered to be a justified public health strategy, and when the Supreme Court heard Buck v. Bell (1927), they ruled the practice constitutional "for the protection and health of the state." Though these practices began inside of institutions, the precedent set by Buck v. Bell allowed states to begin sterilizing individuals that had not been institutionalized, often through coercion and deceit. Minnie and Mary Alice Relf, two Black sisters, were just 12 and 14 years old when they were coercively sterilized. These sterilizations were performed at a Health, Education, and Welfare clinic in Montgomery, Alabama. When the Southern Poverty Law Center filed Relf v. Weinberger (1973), a "district court found an estimated 100,000 to 150,000 poor people [often BIPOC] were sterilized annually under federally-funded programs." This denial of the right to procreate, the limiting of the agency of women of color, and the eugenic goal of these programs represent a major misery and a prevention of being born.

Thus far in this section, I've focused on the descriptive features of health status and outcomes for different racialized groups, showing how racialization is connected in many cases with premature death, worse health outcomes, a lower quality of healthcare, and a variety of other miseries related to health. In addition to these descriptive features reported through data, much of the research that explores health outcomes is focused on *explaining* why racialized groups experience health inequalities. These resources can be used to establish the differential experience of miseries that lead to premature death when experienced throughout a lifetime.

In a (2015) study, Phelan and Link wanted to determine whether or not race[49] was a fundamental contributor to health disparities independently of socioeconomic status (earlier studies [Link and Phelan 1995] concluded that socioeconomic status was the primary fundamental contributor to health inequalities). Through their new analysis, they identified several mechanisms of lower health outcomes that are race-based: experiences of discrimination, other stressors (difficult life circumstances), weathering (cumulative stress), low-quality healthcare, neighborhood conditions, and resource access (recreation, nutrition, harmful substances, protection and crime, toxic environmental exposures, and medical services). These mechanisms are identified as sustaining causes[50] to show why healthcare outcomes are unequal between racialized groups. We can use these features to identify different miseries faced by some racialized

---

[49] This study only looks at differences in Black and white health outcomes.
[50] In contrast to simple determinant causes that lead to explanatory monism.

groups (e.g. the experience of living next to a waste management site) and the detriment these miseries have to individual and collective healths.

As discussed in Phelan and Link (2015) the experience of discrimination also has a negative influence on health. In the context of the United States (and other similarly situated locations), where racialized groups are the targets of discrimination in their everyday lives, the social stress of expecting and experiencing discrimination has a negative impact on health over the course of a life (Paradies 2006; Williams and Mohammed 2009; R. Harris et al. 2012; Lewis, Cogburn, and Williams 2015; Gee et al. 2019).[51] In addition, studies show that children and young adults who experience discrimination already experience decreased mental health (Priest et al. 2013). Thus, the experience of interpersonal racism can be invoked in causal explanation of poor health outcomes and early death.[52]

The final location of miseries and high morbidity for BIPOC I'll discuss is mass incarceration and police violence. Though they make up a only 32% of the population, in 2015 African Americans and Hispanics represent 56% of all incarcerated people (NAACP). Alexander ([2010] 2012) argues compellingly that mass incarceration is the new "colorblind" version of Jim Crow laws that serves to control Black and Brown populations and maintain the racial status quo. Black and Hispanic people are also more likely to be threatened with or experience non-fatal force in a police interaction with 1.4% of white people experiencing force and 1% experiencing excessive force, 3.5% of Black people experiencing force and 2.5% experiencing excessive force, and 2.1% of Hispanic people experiencing force and 1.4% experiencing excessive force (Hyland, Langton, and Davis 2015). Per capita, Black people are much more likely to die from police violence than white people with a major contributor being the higher rates of stops and arrests for the Black population (Miller et al. 2017).

I've reviewed this literature to demonstrate that misery, health, and premature death differs across racialized groups and is concentrated among BIPOC, particularly among Black and

---

[51] These studies are meta-analyses, all of which consider the influence of discrimination or racism on the health outcomes of a variety of oppressed racial groups. Thus, these studies are able to show decreased health outcomes for a variety of oppressed racial groups (in contrast to the study I mentioned above that focused on African Americans).

[52] I'll discuss why this is important through my discussion of condition B below and in Chapter 4. In terms of our individual and collective responsibility, the experience of racism might be the most pervasive factor since the experience of racism is constituted by the many millions of subtle interactions people have throughout a life time, rather than in one primary event (like a heart attack or explicitly hateful interaction, etc.).

Indigenous[53] people. I've used this data to motivate the idea that racism has primarily to do with these features and that looking more closely at the experiences of racialized groups will help us improve our epistemic position with respect to the concept RACISM, and the application conditions for the predicate 'racist'.

3.3.4 That's Racist! Application Conditions for the Predicate 'Racist'

Harris' actuarial account of racism and the empirical data reviewed above guides my selection of the application conditions for the predicate 'racist' throughout the paper. My articulation of the application conditions for the predicate racist grounded in the descriptive harms of racism moves from the "logical structure" way of thinking about racism and attempts to avoid the problems Harris identifies with logical/explanatory accounts. I will move forward from Harris' (2018) position, unjustified impending and premature death as morally wrong, in my argument. As I will describe them, the application conditions are not non-moral (or neutral with respect to moral theories). But, unlike the explanatory or "logic-based" accounts identified by Harris, they do not rely on the cause of death for the source of their wrongness. Specifically, I will argue that the harms of racism that lead toward unjustified premature death, no matter the cause, are a violations of the right not to be killed.

In the following subsections, I will argue that the predicate 'racist' is appropriately applied to an entity when the entity plays some role in an individual suffering harms that prematurely kill and prevent from being born and when those harms are: 1. unjustified, 2. suffered due to race, 3. conditionally probable to be suffered given race in a given context. On this account, the satisfaction of condition 1. makes the harm a wrong, and the satisfaction of Conditions 2. and 3. are what make this wrong racism and not some other kind of wrong. The strength of this account is that it avoids being too narrow (avoids anomalies) and captures the wide variety of cases in which we'd like to hold individuals and collectives accountable. In other words, unlike accounts that limit the application of the predicate 'racist' to specific kinds of entities, this account avoids the inability to capture anomalous cases by providing a descriptive account.

I will move forward in this the following sections by breaking down the application conditions into analyzable units. This will help us maintain clarity about the motivations and

---

[53] There is very little data specifically focused on indigenous health outcomes; however, the low life expectancy can be used as evidence that premature death is concentrated in this population.

reasons for each part of the application conditions. I will use the following analyzable units to argue for the application conditions:

The predicate is appropriately applied to

(A) an entity

(B) when the entity plays some role in an individual suffering harms

(C) that prematurely kill and prevent from being born and

(D) when those harms are:

(D.1) unjustified,

(D.2) suffered due to race,

(D.3) and conditionally probable to be suffered given race in a given context.

3.3.4.1 (A) An entity

For this account, I am understanding (A) "an entity" in a broad manner to capture any kind of thing we may want to evaluate as racist. Here, entities can refer to things we typically think of as entities such as physical objects, individual bodies, organizations, companies, and governments. However, I am also thinking of entities which aren't altogether physical—mental states, social and cultural practices and norms, social schemas and frames of intelligibility, and informal processes— as entities for the purposes of the application conditions. The function of (A) "an entity" in the application conditions is being able to pick out something in the world to which we will or will not apply the predicate 'racist'. Because of entity type pluralism, the application conditions are not sensitive to the type or kind of entity that is being evaluated. There is no kind of thing for which it does not make sense to ask the question, "Is X racist?" The answer for some kinds of things will straightforwardly be, "no." For example, if I ask, "Is the number 7 racist?" My answer will be, "no." But, it has not been a nonsensical question. Thus, on this account, we will be able to ask seriously whether or the Long Island Parkway bridges are racist.

3.3.4.2 (B) when the entity plays some role in an individual suffering harms

The specificity of (B) "when the entity plays some role in an individual suffering harms" is a bit more challenging, especially in practice. One place we might look to identify the relationship between entities and harms is the moral literature on responsibility, specifically attributability, answerability, and accountability (Scanlon 2000; Smith 2007; Shoemaker 2011).

However, a problem immediately arises: responsibility is reserved for moral agents, and many of the entities I'm concerned with in this paper are not moral agents (or even agents at all). So, this route isn't very promising. However, not all hope is lost, we can look to mechanistic explanation in philosophy of science and functionalism in philosophy of mind to help us link entities (coherent bodies of some sort) to individual instances of harm.

In "Thinking about Mechanisms," Machamer, Darden, and Craver (2000) argue that we should understand the practice of biology, and perhaps science as a whole, as a search for mechanisms. Mechanistic accounts include a phenomenon to be explained, the entities that are involved, causings among entities, and the organization of such entities. Bechtel and Abrahamsen (2005, 423), write "A mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena." Understood through this mechanistic lens, feature (B) establishes that the entity identified in feature (A) *has a place* in a mechanistic explanation of the phenomenon in question (something that we want to identify as racist, possibility). Note we do not need to understand or be able to fully explain a mechanism in order to identify that an entity should play a role in a mechanistic explanation.

We can see this even more clearly in Lycan's (1995) role/occupant distinction (38) within functionalist accounts in philosophy of mind. In describing this typical distinction in contemporary functionalism in philosophy of mind, Lycan (1995, 38) writes, "the seductive comparison of people (or their brains) to computing machines drew our attention to the contrast between a machine's program (abstractly viewed) and the particular stuff of which the machine happens to be physically made, that *realizes* the program." When we're going about the business of explaining the behavior of computers (or humans), we're interested in the machine's program, the *role*. We're not so interested in the *occupant*, or the stuff that the program is physically run on. Lycan is articulating this distinction for a particular purpose—to improve our understanding of functionalism—but we can take this distinction to increase our understandings here. When we identify that (A) "an entity" has (B) "played some role in an individual suffering harms" we are saying that (A) *is an occupant* (the hardware) that realizes the role (the software) of racism. We

are primarily concerned here with the picking out the stuff in the world[54] that realizes the "program" of racism. If (A) is determined to be an occupant, then condition (B) is satisfied.

Part of my goal in articulating (A) and (B) in this fashion is to turn up the resolution on our world and look at entities in a far more fine-grained fashion. At first blush, there are many cases where a particular entity plays a clear and primary role in causing harm to an individual, or at least a clear role on its face. Let's start with a well-known case: On February 26, 2012, Trayvon Martin, a Black seventeen-year-old young man, was shot and killed by George Zimmerman, a twenty-eight-year-old man who coordinated the neighborhood watch. George Zimmerman, as an entity, caused harm (in this case that led to immediate death) to Trayvon Martin. George Zimmerman, as a whole person, should be included in our analysis of the case. But, we can also quickly complicate this story by including in our evaluation of the case Zimmerman's mental states, including the schemas he had access to for understanding the interaction, and Florida's laws around gun regulation and self-defense.[55] These entities (Zimmerman's mental states and the gun laws) are occupants that run the program of racism and satisfy (B).

Turning up the resolution on this case sprouts new entities to worry about, but it doesn't call into question that these entities played a role in the case. Remember that my account of the application conditions for the predicate 'racist' doesn't require that the causes are coordinated together in some common logical structure. Thus, in order for us to identify (A) and (B) we just need to be able to articulate the case in a way that confirms that the entity did play a role in the harm.

One worry that arises about my articulation of (B) "when the entity plays some role in an individual suffering harms" is that there are some people with very intuitively racist beliefs who don't end up playing any role in individuals suffering harms. Take Blum's (2007, 9) example:

> …suppose a white person hates Black people, but the white person is relatively isolated and powerless, does not come into contact with Blacks, and generally does not even express his or her feelings to others. This is unquestionable racist, yet it

---

[54] Occupants are also sometimes called fillers, as in the things in the world that fill the roles. Interestingly for our purposes, they are also sometimes called structures, which Lycan (1995, 47) thinks is misleading since, "a *structure* is an organized collection of elements, somehow *held in place* and/or serving to hold other things in place for some purpose or another." On my account, however, structures can be occupants, but this is a different use of the word structure than is typical in the literature on functionalism.

[55] Bringing in other entities, especially to explain why something happens, is often a way to immediately lessen personal responsibility. One might think that because there are other entities that we can identify as playing a role in the death of Trayvon, that this lets Zimmerman off the hook. However, this is not yet a story about blameworthiness (see Chapter 4 for more details).

contributes virtually nothing to a system of unjust advantage based on race. If racism were only systemic, then such an individual would not be racist and, indeed, individuals could never be racist.

Blum's target here are theorists who focus primarily or solely on systemic or institutional racism (structuralists), but this worry also applies to my entity type pluralist view. Thus, perhaps we should revise (B) to "when the entity plays (or would play given the appropriate connection to other entities) some role in an individual suffering harms." In other words, the entity would play a role in some individual suffering harm in some other close possible word given its relevant features. In the case that Blum gives, the relevant features are that content of the beliefs, not the social isolation of the person. Articulating (B) in this way, gets at the worry Blum has identified without needing to give a special account of personal racism that is independent of its effects.[56] However, as will be discussed in Chapter 4, my intervention-senstive account of moral responsibility for racism will be less concerned with these sorts of cases.

In this subsection, I've explored what it means for an entity to be play a role in causing harm. In the next subsection, I'll focus on the primary moral feature of my account of the application conditions for the predicate 'racist': an individual suffering harms that prematurely kill and prevent from being born.

3.3.4.3 (C) (an individual suffering harms) that prematurely kill and prevent from being born

Harris' (2018) account of necro-being remains agnostic with respect to the normative framework that would motivate premature death as morally wrong, assuming any reasonable moral framework would find it wrong. In this section, I'll ground the wrongness of premature death by appealing to O'Neill's (1975) "Lifeboat Earth" in which she famously articulates vast array of conditions under which the violation of the right not to be killed can occur. I use her conception of the violation of the right not to be killed primarily because it helps motivate a notion of *cumulative premature death*, a death that occurs as a result of multiple processes over time.[57] I

---

[56] Another way to go about this is to say that racist beliefs harm the individuals who have them. There's some evidence for this, for example, take those who kill themselves after committing a hate crime. I ultimately want to argue for something like this, but in this chapter, I'll focus on the harms of racism that are experienced by people of color.

[57] Though I take on O'Neill's rights-based framework here, I think we can easily take her notion of cumulative death away from the paper without accepting that rights violations are the best way to identify wrongs in general. So, like Harris, I'd like to remain somewhat agnostic with respect to the frameworks I take on wholesale.

think this notion helps us understand why health, death, and dying is the primary descriptive concern of racism.

In the paper, O'Neill assumes that persons have the right not to be killed and, rather than arguing for it, the distinctions put forth in the paper help us to understand this right, as well as the nature of the wrongs that are inflicted when this right is violated. In addition, from the right not to be killed is a corollary duty not to kill. Her goal in the paper is to begin to understand our moral responsibilities with respect to global hunger and famine and to fend off apathetic and exculpating responses of the "I'm not involved or implicated in these deaths"-type. Most important for my analysis are the conditions O'Neill identifies in which an individual or group violates another's right not to be killed, but when our intuition is to absolve those involved.

O'Neill (1975, 280) argues that "B may violate A's right not to be killed even when:

    1. B does not act alone

    2. A's death is not immediate

    3. It is not certain whether A or another will die in consequence of B's action

    4. B does not intend A's death."

These conditions help us begin to sort out complex questions of right violation and moral failures. The most important insight we take from O'Neill's account is that harms that lead toward death do not always lead to immediate death caused by one person where that one person intended to unequivocally kill the other (though this is sometimes the case). Rather, the most complex cases involve a variety of actors that unintentionally kill others slowly over a long period of time and where the single cause of death is indeterminate. In other words, there are many cases where premature death is *cumulative*. This process is not foreign to us. It is now considered common knowledge that smoking cigarettes kills people. No one argues against this. Further, no one complains that this isn't *really* true because someone's first cigarette doesn't lead to immediate lung failure and death. We already have an intuitive understanding of something contributing to someone's premature death, slowly, cumulatively, over time.

So, what we take away from O'Neill's account is that the right not to be killed can be violated in cases where the causal story for premature death is complicated. That the story isn't "B killed A immediately and on purpose" doesn't matter in evaluating whether or not the right to be killed has been violated. Thus, we can use her account to understand the kind of premature death I'm interested in throughout the chapter.

Now, what about the prevention of being born? The United Nations Convention on the Prevention and Punishment of the Crime of Genocide (1951) made acts of genocide punishable under international law. There are five types of acts that come under the umbrella of genocide, and the later two have to do with children: imposing measures intended to prevent births within the group and forcibly transferring children of the [national, ethnical, racial or religious] group to another group. As mentioned in the miseries section, the United States has a history of coercive or forced sterilization of women of color. As recently as 2010, up to 150 women incarcerated in California prisons were coercively sterilized (Chappell 2013), and since more than half of incarcerated women in California are women of color, many of women of color were likely sterilized. Because birth is essential to a group of people living on and (you might think) that having and raising children is important for the wellbeing of communities, the prevention of being born is a manifestation of misery and the premature death of a people. In addition, though preventable infant fatality is certainly a premature death, it also seems to align with worries within the prevention of being born. Thus, we should include the harms associated with the prevention of being born in our application conditions for the predicate 'racist'.[58]

This will do for our understanding of (C) "suffering harms that prematurely kill and prevent from being born." The purpose of this larger section is to articulate the application conditions for the predicate racist. I argue that the predicate is appropriately applied to (A) an entity (B) when the entity plays some role in (C) an individual suffering harms that kill and prevent from being born and (D) when those harms are: 1. unjustified, 2. suffered due to race, 3. conditionally probable to be suffered given race in a given context. Thus far I've clarified (A), (B), and (C). These three conditions work together to identify entities that occupy the roles within process that lead toward death, but we haven't yet made the connection to racialization. So, I'll now move to (D), the three conditions that such harms must meet in order for the entity playing a role in the harms to be appraised as racist.

---

[58] For the purposes of my discussion here, I will bracket the "abortion as Black genocide" debate. Though I do not wish to deny that Margaret Sanger had racist beliefs and supported some forms of eugenics, I do wish to side with Black women who argue that abortion care is essential for Reproductive Justice (Ross and Solinger 2017). For brief insight into the ways in which anti-abortion advocates have appropriated Black Lives Matter, read https://www.vox.com/identities/2018/1/19/16906928/Black-anti-abortion-movement-yoruba-richen-medical-racism.

3.3.4.4 (D) When those harms are: 1. unjustified, 2. suffered due to race, 3. conditionally probable to be suffered given race in a given context.

Analyzable unit (D) of the application conditions for the predicate 'racist' has a big job to do. (D.1) takes the harms identified in (C) and shows that they are wrongs rather than mere harms (as in, the harms violate some normative requirement). (D.2) and (D.3) identify that the wrong that has occurred has to do with racism and not some other wrong. Further, (D.3) will move us forward in understand the systemic nature of the wrongs of racism as opposed to other interpersonal harms. In the next three subsections, I'll discuss each condition in turn.

3.3.4.4.1 (D.1) When those harms are unjustified

In order to motivate (D.1) "when those harms are unjustified," I'll turn back to O'Neill's account of the violation of the right not to be killed. On O'Neill's account, there are some justified violations of the right not to be killed: unavoidable killings and self-defense. The case of unavoidable killings is most easily understood through philosophical cases in which a person has no choice but to choose between killing two people because they are being forced or controlled in some way. However, O'Neill also uses unavoidable killings to explain deaths in "ill-equipped" lifeboat cases. On this account, the right not to be killed can be justifiably violated when in the actual state of affairs there are not enough resources on the earth for everyone to be supported. At the time of her writing, O'Neill did not think that the earth is ill-equipped with resources. Thus, she asserts in her paper that there are no unavoidable killings given the current resources. Rather, O'Neill argues that most cases in which we may take the problem to be lack of resources are really distribution and hoarding problems.

This will be the case for the harms of racism as well. Harris (2018, 40) writes, "The relationship between dominant and subjected groups is one in which health can be understood as transferred from one to the other." It is not that there isn't enough health or life to go around (i.e. there is no zero-sum game), but that resources and the conditions that support health are directed toward dominant groups, and they are taken from the subjected by the dominant. Death is not unavoidable in these cases.

According to O'Neill, self-defense is the other justified violation of the right not to be killed. Self-defense is slightly more complicated and plays a role in more cases. However, O'Neill's (1975) conception of self-defense does not include psychological features such as

"feeling threatened" or as "though one is in danger." Rather, for the violation of the right not to be killed to be justified, the one who kills' life must be *actually* be in danger.[59] Using her lifeboat metaphor, O'Neill shows that cases of self-defense are limited in a well-equipped lifeboat situation (no unavoidable deaths). For example, if A threatens to jettison the fresh water, and B kills A to save the water and prevent A's killing of everyone on the boat, B justifiably killed A in self-defense (277). However, if B decides to kill A in order to have more water for himself and to ensure his survival, B has not killed A in self-defense. B is thus unjustified in killing A (note again that this avoids the problem of psychological features—like B is *afraid* there isn't enough water—from justifying the killing). Persons' rights not to be killed can only be justifiably overridden in cases of unavoidable killing and killing in self-defense. All other harms that lead toward death are unjustified.

Now that we've discussed O'Neill's account of justified killing (unavoidable killings and self-defense), we can come back with a greater understanding for (D.1) "when those harms are unjustified." (D.1) is met when neither of O'Neill's criteria for justified killing are met. The unjustified condition expresses the morally problematic feature of racist entities, *unjustified* cumulative premature death. However, unjustified harm that leads toward premature death by itself does not specify particularly racial wrongs, just the wrong of violating the right not to be killed. I will now move to Conditions (D.2) and (D.3) to complete my articulation of the predicate 'racist'.

3.3.4.4.2 (D.2) When those harms are suffered due to race

(D.2) is simultaneously the most straightforward and most complicated feature of the application conditions for the predicate 'racist'. In order for an entity to be *racist* (and not some other kind of wrong), it must have something to do with race. So, suffering harms that prematurely kill and prevent from being born must be suffered due to race. However, as soon as we begin to decide whether a harm has been suffered due to race, the questions of what races are and how group membership is determined come to the fore. It is outside the scope of this paper to give a full account of the metaphysics of race and racial category membership, but we'll have to have a working conception in order to understand what it means for a harm to be suffered due to race.

---

[59] Of course, this poses epistemic problems for the life-defender. I take it that these epistemological problems are relatively less bad than accepting that non-factive psychological features (e.g. fear of Black bodies) can justify self-defense.

The account in this paper will assume a realist social constructionist account of races (for an overview, see (Mallon 2006; Blum 2010; Taylor 2013)). On this view, racial category membership is not due to a biological essence that causes innate differences between groups of people.[60] Race, on the other hand, is a socially constructed kind that contingently[61] exists as a group of which one can be a member. Race is modally contingent in that in other possible worlds being white might be very different than it is in this world. The maintenance of the social kind occurs in part because we think about and act as though individuals are members of racial kind groups. This racial thinking influences the material conditions of the world, which further supports such racialized thinking (see Mallon 2016 for more details of an account like this). Social constructionist accounts of race also point to the idea that we do have a sense of racial kind membership because the existence of the kinds in part depends on our mental representations of the racial kinds. Thus, those of us living in a racialized world will have a sense of when harms are suffered due to race.[62]

Now that we have a basic understanding of the racial metaphysics assumed by this account, we can take a look at some examples of (D.2). We'll start with the relatively straightforward examples of harms that result from actions rooted in explicitly racist beliefs. For example, Dylann Roof's murder of nine members of the Emanuel African Methodist Episcopal Church was grounded in his white supremacist beliefs about the superiority and rights of white people and his hatred of People of Color. In this case, Roof himself would agree that the harms suffered due to his actions were suffered due to race because otherwise he would not have committed the murders. Hate crimes motivated by racial hatred are by no means infrequent, and we should take them very seriously as events that satisfy (D.2).

However, there are other harms suffered due to race that are more covert. For example, for many years gynecologists believed that endometriosis was uncommon in Black women (Chatman 1976; Kyama et al. 2004). This led to the misdiagnosis of pelvic pain in Black women as pelvic inflammatory disease, which has a different treatment plan that fails to treat symptoms of

---

[60] This feature is particularly important for showing why differences in life expectancy cannot be explained through appeal to biological features.

[61] See Hacking 2001; Kukla 2000; Haslanger 2012c.

[62] One interesting, and complexity inducing, feature of socially constructed races is that the experience of racism (or race privilege) is at least partially constitutive of being/becoming/continuing to be a member of race. Take, for example, Haslanger's (2012a) account of racialization that includes systematic subordination and privilege. On these accounts, individuals are members of the races that they are members of because of the systematic treatment they receive in a particular context (on Haslanger's account, this treatment is due to and justified real or imagined features of groups).

endometriosis. Fabien (2017) argues that one reason for delayed diagnosis of endometriosis in Black women is that the confirmation test is risky, and gynecologists will not pursue it if they have doubts about the diagnosis. However, the decision to test is based off the testimony of the patient and reports of their pain.[63] Thus, doctors must believe the testimony of Black women in order to pursue diagnosis. A possible and likely explanation here is that doctors prescribe identity-prejudicial credibility deficits (Fricker 2007, 28) to their patients based on race. If this is the case, harm is suffered due to race both epistemically (in the patient not being believed) and materially in her continued suffering due to pain and the potential long-term health effects of misdiagnosis, including increased risk for other diseases such as heart disease and cancer, both of which can cause premature death.

Though more covert than the above case above in that (we at least very strongly hope) doctors do not harbor explicitly racist beliefs, have racial hatred, or intend to harm their patients, this case (and other similar cases of medical bias) is still a relatively straightforward example of an individual suffering harms due to her race, even if the causal story is a bit more complicated (bias on the part of the doctor and medical community, epistemic injustice in distrust of the testimony of patients, long term effects of misdiagnosis, etc.).

We also need to be interested in and aware of cases in which harms (or costs) of a public good are distributed unequally across the population along racial social kind category memberships. For example, when a city is faced with the need to build a waste management site, we might think they are providing a public benefit to the community by supporting commercial endeavors in their area. In the process of choosing a site for the facility, however, they will often concentrate the costs, costs often directly associated with health outcomes, on particular communities due to the physical proximity to the facility. The site may be chosen because the individuals that live there do not have the collective bargaining power (read: capital) to prevent the facility from being built in their neighborhood. This lack of capital is maintained through other racialized systems of resource distribution (e.g. red lining, school funding).[64] Though they receive

---

[63] Racial bias in uptake of reports of pain are well documented. Hoffman (2016) and her colleagues argue that this occurs due to mistaken beliefs about the biological differences between Black and white people.

[64] When we get to this level of causal complexity, Frye's (1983) bird cage analogy is useful. She argues that when you zoom into the level of just one wire, it is difficult to see why the bird cannot escape. However, when one zooms out to the level of the cage, it is immediately clear. This is why the causal stories of oppression are difficult to sort out. Thankfully, this won't be a problem for my account. Remember, what I hope to do in delineating the application conditions is to provide a pluralistic account of the application conditions for the predicate racist. This does not commitment me, and in fact intentionally avoids, to one neat causal story to explain the origin of racial harms.

a very limited amount of the benefits, they bare most of the (health) costs of being close to the facility (Frost et al. 2005). These examples illuminate the kind of thing I mean by (D.2) "a harm being suffered due to race." Taken together, these conditions help us identify the morally problematic features of racism, but we still need to take one more thing into consideration: what about "reverse racism"?

3.3.4.4.3 D.3 When those harms are conditionally probable to be suffered due to race in a given context

One might think that (D.1) and (D.2) are sufficient for delineating the kind of wrong that is racist since they identify the conditions that must be met for a harm to become a wrong and for the wrong to be racist. However, these conditions alone leave room for both "racism against white people" or "reverse racism" in every context. I'm particularly interested in cases when these concepts are invoked when white people are the targets of moral appraisals and accusations of racism (e.g. you're the real racist!)[65] and when they are used to critique policies such as affirmative action. It is my hope that the inclusion of (D.3) "when those harms are conditionally probable to be suffered given race in a given context" within the application conditions keeps the predicate 'racist' from being applied in these specific cases. My primary motivation for this condition are the contingent historical facts around which individuals and groups have been the targets of harms suffered due to race. It would be tempting here to bring in a "proper cause" account of why some things are racist and others are not, even when both things meet conditions D.1 and D.2. But, given my goal of explanatory pluralism, this is not an option available to me. Thus, I bring in two descriptive features: "conditionally probable to be suffered given race" and "in a given context."

"Conditionally probable" here is meant to connect the descriptive account of racism with the patterns of power over death without committing myself to explanatory monism. One might think that cumulative premature death already accounts for these patterns because only someone who is systematically subject to the miseries, healthcare deficiencies, and other harsh conditions will actually die prematurely from such patterns. Thus, even if a white person is the target of discrimination from one shopkeeper or from one person they pass on the street, it isn't going to

---

[65] This also occurs when people of color refuse to engage with specific white people or white people in general because it is harmful to their health (or a waste of their time, energy, etc.). Though this is justified in that it is a measure to protect oneself from damaging effects of racism, this is often labeled as racism by the white people who experience this rejection. It my primary goal in this section to show why appraisals of these types are incorrect.

lead to cumulative premature death because the stress is not together frequent, over a long duration, and "ambiguous, negative, unpredictable, and uncontrollable" (Williams and Mohammed 2009, 32-33), since these are the features of stress that are associated with long-term negative health outcomes.

But, there are cases where racial prejudice on the part of a group of individuals of color does lead to major harms for white people. As discussed in Blum's (2007) chapter "Can Blacks Be Racist?", there are a limited number of cases where this occurs. Out of the 1020 hate groups tracked by the Southern Poverty Law center in 2018, 264[66] were Black Nationalist; however, in these records, there are no reported notable events such as beatings, murders, or lynchings of white people coming from these groups.[67] In addition, the SPL counts the Black Panther party as a hate group, which would clearly be hotly contested. Thus, these groups can be understood as having a political ideology that is politically against sustained white power due to historical conditions.[68] See SPL's explanation of the inclusion of Black Nationalist Hate Groups:

> The racism of a group like the Nation of Islam may be the predictable reaction to white supremacy. But if a white group espoused similar beliefs regarding African Americans and Jews and [sic], few would have trouble describing it as racist and anti-Semitic. If we seek to expose white hate groups, we cannot be in the business of explaining away the Black ones.

Note that the SLP definition of racism is primarily (if not solely) about beliefs or "ideologies," rather than patterns of miseries and premature death. Similarly, Blum argues that People of Color can be racist because of their beliefs and racial hatred independently of the systematic patterns of power.[69] He (2007, 39) writes, "it is the content of attitudes and beliefs that makes them racist, not whether their possessors have the power to put them into practice." So, we find that both SPL and Blum's arguments that people of any race can be racist to stem from their definitions of personal racism that focuses on the racial interiorization or antipathy of others by individuals (Blum 2007, 9). Blum's argument that People of Color can be racist against white people relies on his account

---

[66] See https://www.splcenter.org/hate-map. This also increased from 233 in 2017 and 193 in 2016, which is further evidence that these groups are often a reaction to current political conditions.

[67] The Anti-Defamation League (ADL) claims that there have been fifteen murders by Black Nationalists in the U.S. over the past 20 years ("Murder and Extremism in the United States in 2018" 2018), compared to 39 murders by White Supremacists in 2018 alone.

[68] Like the SPL, I do not wish to "explain away" the ideologies of these groups, especially as it relates to anti-Semitism and anti-LGBT stances. I do think, however, that an ideology against white people is a far different kind of ideology.

[69] I don't wish to ignore the violence that a very limited number of white people endure from people of color. I just want to say that this is a topic for another discussion.

of personal racism. Thus, I include "conditionally probable" in the application conditions to rule out these cases from our considerations here (not that we shouldn't consider them at all).

In addition to "conditionally probable," I've included "in a given context" because the problems of "reverse racism" with which I am primarily concerned occur within a United States and Western European context. Blum (2007) also offers compelling cases in which racial hatred occurs between racialized ethnic groups in other contexts, and I do not want to ignore these cases (e.g. Japanese prejudice against ethnic Koreans living in Japan). In fact, in locations where the majority of the population are People of Color, these may be the primary and most frequent cases. Thus, the statistical likelihood of harms that lead toward premature death should be evaluated from a specific context.

My articulation of (D.3) does not rule out as a matter of necessity that People of Color can be racist, but rather it limits the application of the predicate in a way that is sensitive to the social and political realities in particular contexts. Within the context of the United States, the kinds of wrongs with which I'm concerned are concentrated in populations of People of Color and not in those who have been racialized as white, and I've thus articulated the application conditions in a way that reflects this reality without ignoring the contingency of this concentration. In another world that unfolded in a different way, the predicate racist might capture a different set of realities.

3.3.5   The Connection to Well-Being

A potential reaction to Harris' account of necro-being, as well as the application conditions for the predicate 'racist' that I've built from it, is something like: "I like this account, but *death*? *Death* can't be what's wrong with racism. And, even if it is, I haven't played a role in someone's *death*." Less reactionary are concerns that there seem to be many clearly racist things that don't lead to pre-mature death, but only decreases in well-being. In this section, I'll discuss the connection between well-being and pre-mature death. Specifically, I'll explore the "metaphorical" use of death in which someone's well-being is so damaged that they experience a "living death." Both interpretations of the link between well-being and pre-mature death are useful for parsing out the details of the application conditions for the predicate 'racist'.

As discussed in the miseries subsection of this chapter, lower levels of health and wellbeing are concentrated within populations of People of Color. Intuitively, lower levels of health are intertwined with pre-mature death, but they also constitute the lived misery and tragedy of racism.

Harris (2018, 40) writes, "It [racism] effects [sic] the preconditions for the possibility of embodied wellbeing…The relationship between dominant and subjected groups is one in which health can be understood as transferred from one to the other." The connection between Harris' account and accounts of "living death" is clear through his use of Mbembé's (2003, 11) account of necropolitics in which the exercised power of political systems is the power to "dictate who may live and who must die." Related to my articulation of (C) "(an individual suffering harms) that prematurely kill and prevent from being born" as the violation of the right not to be killed, Mbembé argues sovereignty is primarily expressed through the right to kill. Using slavery as an example, Mbembé shows that one can be kept alive (in a state of injury), but that this life is a living death or a form of "death-in-life" (21). Thus, my articulation of (C) can be interpreted through this living death lens. Indeed, the empirical literature on the experience of racism (discussed in the miseries section above), particularly the vigilance that is required to anticipate and stay alert to incoming threats to ones livelihood, leads to decreased wellbeing and chronic health conditions, such as hypertension and depression.

Living death can also be analyzed through a capabilities approach and Nussbaum's (1995, 81) two thresholds. After identifying several capabilities essential to human life, Nussbaum argues that there are two thresholds of capability to function as a human. The first threshold is the one in which "a life will be so impoverished that it will not be human at all" and the second in which "characteristic functions are available in such a reduced way that, though we may judge the form of life a human one, we will not think it a *good* human life." Both thresholds of human capabilities are relevant to us, but particularly relevant is the idea that a human could be so depraved that we would no longer call the life one is living a human life. A full exploration of the capabilities and the ways in which they are limited through the wrongs identified through the application conditions of the predicate 'racist' would be enlightening, but for now the primary purpose is to show that there are several ways to conceptualize living death.

It is my hope that these brief explorations show that (C) can be broadly interpreted to include forms of living death that result from the cumulative suffering of harms. Though it does shed light on the objection I'm interested in—that *death* can't be what one is implicated in when their action, for example, has been evaluated as racist—it doesn't provide the objector from much relief since living death, though lacking the finality of pre-mature death, is just as morally bad.

3.4   Conclusion

3.4.1   Is that racist? Back to the Bridges

I started the previous section with the case of the bridges over the Long Island Parkway, which are too low to allow buses to pass underneath them, to highlight the ways in which entity type pluralism have influenced my account of the application conditions for the predicate 'racist'. So now we ask, is are the bridges racist? Remember, the predicate is appropriately applied to (A) an entity (B) when the entity plays some role in an individual suffering harms (C) that prematurely kill and prevent from being born and (D) when those harms are: (D.1) unjustified, (D.2) suffered due to race, and (D.3) conditionally probable to be suffered given race in a given context.

We've identified the entities as the bridges, so we've done the work of (A) already. The role (B) the physical arrangement of the bridges play is that they prevent buses, a primary mode of transportation for People of Color, from driving along the only roads to Rockaway Island, a recreation site close to the urban space of New York City. Thus, (C) the bridges harm individuals by placing a major barrier to enjoying recreation (lack of access to recreation contributes to deteriorated health (Phelan and Link 2015)). (D.1) These harms are not caused due to a lack of resources nor through self-defense. Further, (D.2), these harms are suffered due to race given that People of Color are the majority of bus riders and (D.3) are conditionally probable to be suffered due to race given that the population of bus riders is concentrated in populations of color. Thus, the bridges are racist.

I hope that this demonstration of the application conditions for the predicate racist demonstrate how they can be used to identify the morally relevant features with respect to racism and show how entity type pluralism manifests through them.

3.4.2   Disagreement and Epistemic Engineering

Due to the epistemic nature of the project, disagreements about what does or what does not count as *racist* will arise between common usage, the account I give here, and other existing accounts. In some cases, these disagreements may be disagreements about terminology (how and which predicates express the appropriate concepts), and other cases differences in terminology indicate deeper conceptual disagreements (what properties exist in the word for concepts to lock

onto).[70] Blum (2002, 203), for example, argues that we ought develop a wider range of vocabulary for identifying various racial ills, including 'racial insensitivity' and 'racial exclusion', and that we should reserve the term 'racism' for only those things that are related to racist beliefs and "continuous with its [the term's] historical usage." He argues that non-racist racial ills are morally problematic, but he is clear that they are problematic in a different way than racism. As will be clear as I give my own account in what follows, Blum and I are not merely using different words to identify the same world-stuff. We genuinely disagree about the properties that the concepts pick out in the world.

If we disagree, then one or more of us must be wrong, and, thus, we should be able to identify a method for figuring out which of us needs to head back to the drawing board. Because it is neither Blum, my, nor others working on racism's project to merely figure out how people already employ the concept, it won't be enough to survey people and find the most common usage. So, if we can't look to ordinary usage, where should we look? On my view, we should look to race-based suffering in the world to see what features this suffering shares. We run into a whole host of metaphysical and epistemological issues here, issues that are certainly not settled and are outside of the scope of this chapter. Though these questions may still remain, I hope to have offered a compelling set of application conditions for the predicate 'racist' that get us closer to getting the world right (in this case, getting it right means that the application conditions for the predicate perfectly capturing the extension of the predicate) and useful as we attempt to improve our moral lives. Knowing more about our concepts and the properties they lock onto improves our ability to pick out genuine instances of entities with the property *racist*. Our abilities to get it right with respect to moral concepts are the preconditions to our abilities to act ethically.

---

[70] Still more disagreements will arise about the complex relations between predicates, concepts, and properties.

CHAPTER 4.   SO, THAT'S RACIST? OBLIQUE BLAME AND INTERVENTION SENSITIVE MORAL RESPONSIBILITY

4.1   Talking About Racism

In my experiences as a white woman,[71] there are four common reactions to my having identified something as racist. One: nods, snaps, murmurings of "mm-hmm," and other expressions of agreement. Two: uncomfortable silence. Three: "Eh, is that *really* racist?" Four: outrage mixed with vehement disagreement. These are reactions that I don't receive when I claim that something has the property of being blue or of being an agent. This is because 'racist' is a moral evaluation—a *negative* moral evaluation—not merely a description. Thus, 'racist' serves both descriptive and evaluative functions. As Scanlon (1986, 170) writes, "This [a moral evaluation] is description, but given that most people care about the justifiability of their actions to others, it is not mere description." Another way to think about this is through the lens of thin and thick moral concepts (Williams 1985). When 'racist' functions as a moral evaluation, it does so because it is a thick moral concept. Moving forward from my descriptive of the predicate 'racist', I must develop an account of 'racist' as an evaluation because of these features.

As noted by the different reactions I noted, the evaluation as racist receives mixed uptake. For a limited number people—often those involved in anti-racist activism—being evaluated as racist is a welcome tool for accountability (this is not to say that it is pleasant; I don't welcome accountability like I welcome strawberry shortcake). For very many people, however, evaluation as racist is something they would very much like to avoid. Being called racist is one of the worst things they can imagine. These two groups of people, as well as everyone in between, have different understandings of what an evaluation as racist means for them. Philosophical work can help us clarify these understandings so that the moral function of the predicate 'racist' can be preserved, and we can hold each other accountable toward anti-racist aims.

A recent experience illustrates the complexities of people's reactions. I was at a bar with some friends, and we were talking about racism. At some point in the conversation, a white man

---

[71] My experiences as a white woman describing something as racist or "calling out" racism often differ greatly from the experiences of People of Color who describe something as racist or "call out" racism. For example, I will likely never get accused of pulling the "race card" or being overly sensitive. This is a facet of race or white privilege, which I have not discussed fully in this dissertation, but take as a given (see McIntosh (1988) and Yancy (2008) for more insight into race privilege).

approached me and aggressively asked, "So, you think all white people are racist, huh?" Of course, these situations have their own unique features that make them difficult to navigate—we don't know each other, we haven't built any trust, we're in a bar, our conversation is very public—but there's something that makes this question even more difficult: depending on exactly what he means, my answer changes. Though left unarticulated, I think there are three objections hidden in his question, and in my position as a white person wanting to do my part to have conversations about race with other people, I think I need to take these objections seriously for two reasons. One, it's good that these conversations are happening, and I want to support serious inquiry into the responsibilities white people have to take accountability for racism. And, two, as suggested by the goal of increasing clarity our racism, I think there are some confusions that keep us from collectively having productive conversations. So, here's what I take his three objections to be, the third of which is the motivation for this chapter.

First, he's objecting to being labeled a racist. As I've discussed throughout the dissertation, I think one fruitful pathway is to be more specific about the entities to which we apply the predicate 'racist'. When we do this, we avoid the problems that arise from labeling whole persons as racist (with the exception for avowed racists, white supremacists, Nazis, those who commit racial hate crimes, and others with similar combinations of mental states, dispositions, and actions).[72] Second, I think he means to object against some kind of biological essentialism or determinism about whiteness. I also reject this view, so I won't discuss it here.

Third, is an unarticulated objection to being blamed. Specifically, it is an objection to being *straightforwardly* blamed. Straightforward blame is the kind of blame with which we are familiar: someone has acted or failed to act in a way that violates a normative standard, and we behave in ways toward them that indicates that we're unsatisfied with their behavior. In other words, they've done something wrong, and they're being held to account for it. Straightforward blame captures instances were individuals are blamed in virtue of φ-ing, taking some action.

However, this kind of straightforward blame doesn't capture all the sense of blame we need in order to understand responsibility for racism. I'm sure the reader is familiar with objections like, "I didn't own slaves, so I do not need to pay reparations." This objection is an objection to

---

[72] Another way to make evaluations about a whole person is with the concept of a "mosaic self" developed by Sripada (2016). On this view, who a person *really is* is made up of a complex set of principles, cares, and commitments that can be in conflict. On the view I've articulated, it makes sense to think of a whole person as racist when the fundamental pieces of their identity are racist.

straightforward blame: "I didn't φ, so I don't deserve to be blamed." In order to more adeptly handle these cases, I will introduce the concept of *oblique* blame, as a contrast to *straightforward* blame. This kind of blame allows us to understanding cases of blame that are not in response to a normative violation but are rather invitations for individuals to take anti-racist action. We need this concept of blame to answer the drunk man's objection.[73]

In Section 4.2, I situate my strategy in this paper against Blum's (2014) strategy of avoiding "conceptual inflation" of RACISM. This section highlights how this paper is situated within the arguments of Chapters 2 and 3. In particular, I argue that communicative difficulties within evaluation, like the ones described above, should not influence questions of moral status, of what is and isn't racist.

In Section 4.3, I describe straightforward blame in order to lay out a distinctive kind of blame, oblique blame, that describes blame directed at agents in their relationship to non-agential entities with negative moral properties. When an agent is obliquely blamed, agents are blamed at an angle, not for something they did or didn't do, but to motivate them to take action. For example, I might direct oblique blame at my newly elected Senator because of her relationship to a racist state law, due to her new acquired capacity to intervene on that law. My blame may take the form of calling her to tell her that I'm troubled by the law, organizing others to do the same, and engaging in citizen lobbying. After describing oblique blame, I discuss the virtues of oblique blame in comparison to straightforward blame for addressing the complicated matrices of responsibility for racism. From my conception of oblique blame, I develop an intervention-sensitive sense of moral responsibility, which captures the position individuals are in to intervene on entities with the negative moral property *racist* (as well as other bad social properties).

In Section 4.4, I situate my account of oblique blame and intervention-sensitive responsibility within the dominant literature on moral responsibility, starting with Strawson and moving through three ideas developed by his successors. Specifically, I discuss oblique blame with respect to the method of looking toward "ordinary practice," the ethics of blame, and attributability requirements on moral responsibility. It is my hope that exploring these relations will foreground

---

[73] To be clear, I'm not arguing that people can never be straightforwardly blamed for their racist actions or omissions. This is also common. My goal in this chapter is to introduce this other kind of blame, oblique blame, that accounts for the times in which individuals are blamed at an angle in their capacity as someone who can intervene on non-agential racist entities. In addition, someone can be straightforwardly blamed for consistent unresponsiveness to oblique blame.

any questions that may arise for the reader. In Section 4.5, I conclude with some final thoughts about oblique blame and pave the way for my discussion of interventions into epistemic norms in Chapter 5.

4.2   Evaluation and Communication

The strategy of introducing a new kind of blame into our understanding avoids the problem of allowing complications with evaluation to influence questions of moral status. As discussed in previous chapters, Blum (2007, 2002, 2014) is concerned about conceptual inflation of the concept RACISM and the resulting inability for "racism" to be used as an effective tool for accountability. His primary concern seems to be that people mean different things when they discuss racism. He (2014, 60) writes:

> If Jane uses "racism" to refer to anything that can go wrong, racially speaking, while Lourdes uses it more narrowly for egregious wrongs, Lourdes will feel that Jane is morally overloading behavior, attitudes, and so on that are lesser faults, while Jane will feel that Lourdes is failing to acknowledge racism.

With this discussion, Blum highlights the communicative failures that can occur when we evaluate entities (especially persons) as racist or call one another out with respect to racism. When we mean different things by 'racism', we are unable to effectively communicate. As discussed previously, Blum's strategy for addressing these failures is to reserve the term racism for egregious wrongs caused by an individual's racial hatred or antipathy. His account limits racism by constraining his account by entity type (attitudes and beliefs). However, as I've argued in the previous two chapters, this solution is untenable. So, we need a new strategy for dealing with communicative failures like the one between Jane and Lourdes.

In the previous chapter, I argued for a set of application conditions for the predicate 'racist', arguing that these application conditions capture the extension of the predicate 'racist' and improve our epistemic position with respect to the concepts RACIST and RACISM. In this chapter, I'll explore the role 'racist' has to pay with respect to moral responsibility and blame. I'll move from the *descriptive* projects of Chapters 2 and 3to build the *evaluative* component of this project. In other words, in Chapters 2 and 3, I set out entity type pluralism and the application conditions for the predicate 'racist' in order to improve our epistemic position, to increase clarity, with respect to the concept RACISM. In this chapter, I work from these understandings in order to preserve the

moral function of the predicate 'racist' so that we can hold each other accountable on individual and institutional levels.

As the reader will remember from Chapter 2, *evaluative* projects focus primarily on the kinds of entities involved in a particular moral wrong and the nature of the entities involved. They may ask questions like: What kind of thing is involved in the wrong? Does the entity involved with the wrong have volition or agency? If so, can it/they do something about stopping the wrong? If not, what can we (entities with agency) do about it? What should we do about it? These are important questions, especially if we are interested in preserving the moral function of the predicate in order to effectively hold each other accountable. Answering these questions is the goal of this chapter.

However, as my account of *nested questions* requires, I will build my evaluative account in such a way that is sensitive to the dependence relationship between the questions that must be answered within accounts of racism. Specifically, I will not let answers to questions of *evaluation* constrain my answers to the descriptive questions of moral status. So, even when evaluation and practices of accountability are difficult, I will rely on my account of the application conditions for the predicate 'racist' to tell me when and why we should apply the predicate. For example, it may be that applying 'racist' to a social practice will cause some people to hang on more tightly to the social practice, the exact opposite of the desired effect. Though we should be sensitive to this fact in our practices of holding accountable, features such as these will not affect the metaphysical feature of the entity having the property *racist*. This chapter provides a pathway for action after the *descriptive* determination has been made. What do we do after we identify that something is racist?

4.3 Oblique Blame

4.3.1 Desiderata for an Account of Moral Responsibility for Racism

In this section I will develop an account of a distinctive type of blame[74] I call oblique blame that stands in contrast to straightforward blame. I will then use oblique blame to develop an account

---

[74] Many accounts of moral responsibility and blame attempt to provide a unified account of both. In other words, they attempt to identify the features that are similar to all instances of moral responsibility or those features that are constitutive of moral responsibility. Those who are committed to presenting a unified theory but do not think that all

of intervention-sensitive moral responsibility. The primary feature of oblique blame is that an agent (T) is blamed by another agent (S) in virtue of an agent's relationship to a non-agential entity (E) that has a negative moral property *rather than* being blamed by S in virtue of φ-ing. In other words, the agent is blamed *obliquely* or at an angle. This account of blame, and the further account of intervention-sensitive moral responsibility, satisfies several goals I have in developing an account of moral responsibility for racism and other forms of oppression.

Figure 4 Straightforward versus Oblique Blame

1. The account I present in this chapter accords with and answers questions that arise for the entity type pluralist account I gave in Chapters 2 and 3. On this account, non-agential entities (mental states, cultural practices, social meanings, policies, laws, and institutions) can all be racist. But, non-agential entities cannot be blameworthy, nor are they apt candidates for blame. Thus, if we want these non-agential entities to change, this raises questions about how individuals can be held responsible for such racist entities. This chapter helps us understand the relationship between agents and racist non-agential entities, and the nature of the responsibility that agents have to intervene on racist non-agential entities.
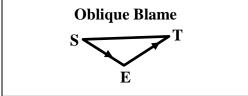2. My conclusions from the Chapters 2 and 3 show what is wrong with the individualist versus structuralist debate. The account in this chapter moves us forward in understanding what to make of responsibility without this debate. In particular, it turns up

---

instances of moral responsibly have shared constitutive features may be functionalists about moral responsibility and blame in that they think all instances of blame, for example, share in their function but not necessarily in form or content. (e.g. Smith 2013). Others, like Shoemaker (2011) insist that there are distinct types of moral responsibility such that it seems giving a unified account is a fruitless endeavor (though he does think that one can give a *comprehensive* account). In presenting a distinctive type of blame, one that differs greatly from dominant accounts, I lean toward an un-unified theory of moral responsibility. I don't rule out the possibility that someone could identify a unified theory that accounts for oblique moral responsibility. I think it is more likely, however, that someone aiming for a unified theory of moral responsibility would reject intervention-sensitive moral responsibility and oblique blame as kinds of responsibility and blame practices because it differs from ordinary practices. However, as I hope to show, we need oblique blame to explain and motivate accountability for racism.

the resolution on how we begin to make change when social forces and constraints are difficult to parse.[75]

3. On the account of oblique blame I develop in this chapter, agents are blamed at an angle, which means that the blame is not an assault on their character or personhood. This helps us further resolve worries that the predicate 'racist' should be reserved for only the worst kind of people since it signals that we no longer want such a person in our moral community (Blum 2002). In addition, oblique blame helps with the kinds of defensive reactions and objections I described in the introduction.

4. In focusing on responsibilities to intervene that are independent of what the agents has done or not done in the past, oblique responsibility avoids problems of attributability. On many accounts of straightforward blame, an agent is only deserving of blame when an action is attributable to them. Much of the literature on moral responsibility has been devoted to identifying which features on the basis of which an action is attributable to an agent. That oblique blame avoids attributability is good because the account doesn't require a theory of attributability and it prevents people from engaging in strategies that allow them to avoid responsibility for racism based on non-attributability arguments (see Levy (2017) for a notable example). In addition, my theory preserves the motivating function of blame without attributability.

5. Oblique blame is agnostic with respect to the emotions the blamer believes the blamed should have or the emotions that the blamer is attempting to bring about in the blamed through their reactive attitudes. Carlsson (2017) has argued that to be blameworthy (in the straightforward sense) is to be worthy of suffering feelings of guilt. However, emotions like guilt, as well as shame and fear, often keep people from taking action toward an anti-racist future (Linder 2015). Since these emotions are often roadblocks to sustainable action, that oblique blame doesn't require these particular emotions is a virtue.

6. Taken together, my account of oblique blame and intervention-sensitive moral responsibility simultaneously raises and lowers the bar with respect to racism. It raises the bar in that people will be called to action far more frequently than they would be with just the concept of straightforward blame. Instances when it will be appropriate to obliquely blame someone will be far and wide. However, this account also lowers the bar in that the concept allows us to talk about responsibility for racism with believing that doing so is a damning expression of one's personal character or the worst thing that one could be called. Intervention-sensitive responsibility is rather a robust acknowledgement of the responsibilities we inherit from living in a world such as ours.

Now that I've highlighted some of the *desiderata* for this project, I'll move forward by reviewing straightforward blame, and then use a case analysis to show how oblique blame differs from straightforward blame. Specifically, it will highlight the ways in which agents can be called to action to intervene on non-agential entities. Unlike straightforward blame, oblique blame says something about what you can do to make a difference in the future, not about how you've behaved

---

[75] I take it that Madva's (2016) paper has a similar goal in developing an anti-anti-individualism account.

badly in the past. I'll use interpersonal cases to highlight the differences, but one can also be obliquely blamed in non-interpersonal contexts.

4.3.2   Straightforward Blame

The moral responsibility literature within philosophy is vast and full of subtleties. For this reason, it is difficult to give an account of straightforward blame that captures every nuance that has been articulated. However, Smith's (2007) rational relations view, as well as Shoemaker's (2011) discussion of her view, is a good place to start since it is a touch point for many theories of moral responsibility. Here, Smith (2007) helpfully separates three senses of what it means for "S to hold T responsible" into T being morally responsible, T being blameworthy, and A blaming T.

On this account (Smith 2007), T is *morally responsible* when his action or attitude (φ) is appropriately related to his evaluative judgments. This means that T could *answer for*, or give reasons as to why, he φ-ed. This is Smith's account of attributability, which makes it one of the more contested parts of her theory of moral responsibility because many of the arguments in the literature center around to what attributability amounts. But, the general point stands that most theories of moral responsibility (and therefore straightforward blame) will give some account of how *actions* (mental or whole-bodied) are related to *agents* in a way that makes agents morally responsible for those actions.

This first sense—T is morally responsible for φ—only captures that an action "belongs" to an individual and would be open for evaluation for the action. So, we need something further. T is *blameworthy* when his having φ-ed violates some normative standard. For many views, the normative standards that matter for blameworthiness are our legitimate demands on one another (Scanlon 2000). This captures the difference between actions that do not violate legitimate demands and those that do. So, it could be the case that T is morally responsible for φ-ing but not blameworthy when φ-ing did not violate a normative standard.

Finally, what it means for S to blame T is for S to judge T as morally responsible, blameworthy, and to direct an appraisal or response [76] toward T that communicates these

---

[76] Throughout the paper, I'll use 'appraisal' and 'response' to talk about the reaction of the blamer (S), in line with Scanlon's (2008) revision of Strawsonian blame. The purpose here is to account for cases in which the limited number

judgements and calls T to apologize and do otherwise in the future. This is what I have been calling straightforward blame. Notice that this notion of blame includes the moral responsibility and blameworthy conditions such that it doesn't include any notion of blame that isn't about *actions* (or omissions, someone could be blameworthy for not φ-ing if not acting violates a normative standard).

> Imagine this case of straightforward blame:

> *Grocery Store:* Kate, a Black woman, is standing in line at the grocery store. Her phone rings, and as she looks down to answer, Jacob, a white man, slips in front of her. She looks up, noticing that Jacob has gone in front of her. Immediately, she makes the facial expression and hand motions that communicate, "Really? Are you serious?" Jacob recognizes that he has been blamed for going in front of Kate in the line and responds by subtly returning to his place in line.

In this case, Kate directly and non-verbally blames Jacob for going ahead of her in the line and violating the norm of fairness with respect to lines. Because of Kate's appraisal (her facial expression and hand motions), Jacob both recognizes and is motivated to repair his violation of Kate's legitimate demand that he not go out of turn. Because this is a case of straightforward blame, this explanation suffices and doesn't leave much of a mystery. Though we could analyze this case more specifically by identifying particular entities (e.g. Jacob's mental states) involved, we wouldn't benefit much from the analysis since the interchange is fairly straightforward and easily resolved. There is no confusion that begs for further clarification. This case demonstrates straightforward blame.

4.3.3 Oblique Blame

Though exchanges like the one above happen very consistently, they don't represent the full range of possibilities when living in a complex world such as ours. Let's take another case:

> *Racist Grocery Store:* Kate and Jacob are both standing in line at the grocery store. The store is very small and busy. The cramped space makes the line ambiguous. Jacob knows Kate arrived to the line before him. When the next cashier is available, the cashier points at Jacob, signaling to him that it is his turn. Jacob has a commitment to undermining the effects of implicit bias and knows that cashiers often call on white people first due to racist implicit biases.[77] Jacob thinks to

---

of moral emotions that Strawson identifies do not come up or would be inappropriate or ineffective. However, there can be cases where the Strawsonian reactive attitudes do come into play, depending on the severity of the case.

[77] I stipulate that the attitudes of the cashiers have the property *racist* and thus the application conditions for the use of the predicate 'racist' are met.

himself that Kate was first, directs his commitment of undermining bias at himself, and motions to Kate that she should utilize the cashier.
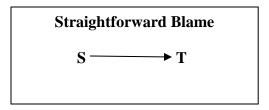
In this case, first, Jacob recognizes that the cashier has a racist implicit attitude that is influencing their actions. Let's say that the cashier's actions are racist as well, not because they result from racist implicit attitudes but because the actions also have the property of being racist. Second, Jacob judges himself to be related to the attitudes and actions of the cashier in a way that allows him to intervene on their influence, at least to some degree, in this moment. More specifically, Jacob knows that implicit biases can have an automatic, unreflective influence on behavior, and because he knows what would be fair in this situation—Kate getting to go first because she arrived to the line first—he is able to intervene on the influence the cashier's implicit biases have on their behavior. Third, Jacob calls upon his prior commitment to motivate himself to act. Perhaps after he leaves, Jacob continues to feel strongly and writes a letter to the grocery store manager about the line management practices, citing the unintended influence of implicit biases on CV review (Bertrand, Chugh, and Mullainathan 2005) as evidence that they need a new "anonymous" system for choosing who goes next in line (e.g. pull numbers).

Racist Grocery Store differs from *Grocery Store* in that Jacob directs a response of conviction toward himself in reaction to a non-agential entity's (the cashier's bias and actions) being racist and then acts to intervene. Superficially, the *Racist Grocery Store* case differs from the *Grocery Store* case in that the blame is self-directed: Jacob directs a response toward himself. But this isn't the feature that makes it distinctive. Rather, the case is distinctive because Jacob blames himself in his relationship to a non-agential entity in a way that motivates him to act. He *obliquely* blames himself. From this case, we can formalize a notion of oblique blame.

> Oblique Blame: For any agent A, any potential agential target T[78] (where T may or may not be identical to A), any entity E, and any negative moral property P: A obliquely blames T if and only if (1) A accepts the evaluation of some E as having P, (2) A judges that T is related to E (but not identical to E) such that T's intervention on E may be effective, and (3) A calls T to intervene on E by directing a response or appraisal toward T.[79]

---

[78] For the purposes of this chapter, I'm thinking of the agential target T as an individual whole person, as these are often taken to be the most straightforward entities with agency. However, in the future, I'd like to develop this account to explain collective agential targets, like governments or other institutions. As Ross (2002) as adeptly argued, humans and collectives share a very similar distributed control system such that collectives are atypical agents only if humans are as well.

[79] My formulation of this definition is inspired by Portmore's (manuscript) definition of comprehensive blame. In particular, I use his language of "potential target T" and "where T may or not be identical to S" in order to adequately include self-directed blame.

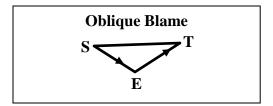| Straightforward Blame	Oblique Blame
S ———→ T	S — T, E

Figure 5 Straightforward versus Oblique Blame

This conception of blame allows for an agent to be held responsible in their relationship to a non-agential entity that has a negative moral property, such as the property *racist*. This is a distinct kind of blame primarily because it does not hold A responsible in virtue of φ-ing (taking some action), but rather in virtue of A's ability to intervene on an entity with a morally bad property. In this way, oblique blame is a call to action, not in virtue of what one has or hasn't done in the past, but in virtue of what one has the capacity to do in the feature. It neatly and effectively avoids objection, "*I* didn't do anything," and articulates the ways in which as individuals we are collectively responsible for dismantling the mechanisms that lead to racist outcomes.

4.3.4   The Virtues of Oblique Blame

One might object that the *Racist Grocery Store* case can be explained by appealing to the normative notions within straightforward blame. One could argue that Jacob would have been blameworthy for following the cashier's directions because his doing so would have violated a legitimate demand that Kate has on him to enforce the rules of fairness. In other words, if he would not have proactively intervened (as is stipulated in the case), Kate could have justifiably blamed him for his failure to act. However, oblique blame is better than straightforward blame in capturing the contours of this case for two reasons. First, within the framework of straightforward blame, Jacob's motivation for upholding fairness is to avoid sanction from Kate, but within the framework of oblique blame, he acts out of his commitment to intervening on implicit biases. Oblique blame better tracks and cultivates anti-racist commitments and motivations to intervene when it is effective.[80] Second, if Jacob goes ahead, and is blamed by Kate, this puts Kate into an awkward position given the social meanings and power relations in place. If she is to hold Jacob to account (or the cashier), she must express a negative appraisal toward either Jacob or the cashier, and this

---

[80] My articulation of the moral wrong of racism being a violation of the right not to be killed underpins the obligation we have to intervene. Oblique blame is meant to capture to mechanism of moral motivation that makes these obligations realizable.

communication could come at a great social cost to Kate. Thus, we need a concept of blame that doesn't force Kate to incur these costs to uphold her legitimate demands on Jacob's behavior.

I think these reasons alone justify entering oblique blame into our moral understandings, but the differences and virtues of oblique blame don't end there. Oblique blame also differs in the emotions that are deserved by T and the requirement that oblique blame be expressed to T (rather than withheld). On some accounts of blame, blaming is said to involve representing T as someone who is deserving of feeling guilt, regret, or remorse (Carlsson 2017, Portmore manuscript). But, in cases of oblique blame, A directs a response or appraisal toward T in order to motivate T to act. In other words, oblique blame is specifically a call to action, a call for intervention. It does not require that S thinks that T is deserving of any *particular* emotion, but rather of any emotion that would motivate T to act. These can include not only guilt or remorse, but might also include vindication, urgency, compassion, or resolve. Oblique blame is forced on the call to action and is agnostic about the appropriate motivating emotions.

This agnosticism is a particularly important piece in articulating an account of moral responsibility and blame for racism since white fragility, which includes emotions like guilt, anger, and fear (DiAngelo 2018), often keeps white people for acting toward an anti-racist future (at least when these are the *only* emotions felt). Linder (2015) finds the same thing. In her analysis of feminist anti-racist identity development, she found that guilt, shame, and fear kept her participants from standing in solidarity with People of Color. Linder (2015, 543) writes, "While participants acknowledged the importance of working through guilt and shame so they could engage in action related to antiracism, they often struggled to do so, creating a cycle between engaging in activism and inaction.". Finally, in my experiences in engaging in anti-racist organizing, the call to action is far more important than feeling any sort of "deserved feelings." In fact, many anti-racist organizing spaces focus on the creation and maintenance of boundaries and white resiliency (Hozumi 2017), so that people can move from these feelings into solidarity.

Oblique blame differs from straightforward blame in that A need not think that T is deserving of particular emotions. An additional difference is about the expression of blame. There are some accounts of straightforward blame that hold that not all blame is expressed (one can keep it to oneself), and as such our definitions shouldn't require that an action be taken (Sher 2006; Scanlon 2008; Smith 2013; Portmore manuscript, 5). However, because my account of oblique blame is an account of how blame might function as a tool for accountability for racism, oblique

blame includes that S take the action of directing a response or appraisal toward T (remember, A and T can be identical, so this accounts for blame that is both self- and other-directed). This is required because without the direction of a response or appraisal, there is no call to action, and the call to action is the primary feature of oblique blame. Thus, there is no private oblique blame. For these reasons, oblique blame is a distinctive kind of blame that is an effective pathway for holding each other accountable for racism in its various forms. Now that I've articulated oblique blame, we can add to it a conception of moral responsibility.

4.3.5 Intervention-Sensitive Moral Responsibility

So far, I've developed an account of what it means to obliquely blame someone. Strawson ([1962] 1993) argues that we should build our accounts of what it means to be responsible from an analysis of blame. In line with this insight, I'll develop an account of what it means for individuals to be responsible for the negative properties of non-agential entities that is derived from my account of holding individuals responsible for the negative moral properties of non-agential entities. I'll call this intervention-sensitive moral responsibility.

> Intervention-Sensitive Moral Responsibility: For any subject/target T, any entity E, and any negative moral property P: T should intervene on E just in case (1) E has P and (2) T is related to E in a way such that T's intervention may be effective.

Of course, this differs from dominant accounts of moral responsibility in the literature since most accounts of moral responsibility are accounts of responsibility for or not φ-ing. In contrast, this is an account about what it means for an agent that is related to a non-agential entity with a negative moral property to be responsible for intervening on that entity.

On this account, moral responsibility and blameworthy-ness are intimately tied together. On other accounts, such as Smith's (2007), to be blameworthy is to be morally responsible and to have violated some normative principle by φ-ing. Since intervention-sensitive moral responsibility focuses on the negative moral properties of non-agential entities, this additional step is not needed.

This articulation of moral responsibility helps us understand individuals' relationships to many different entities—their own mental states, cultural practices, laws, institutions—and the responsibility they have to intervene on them once the negative property has been identified. However, the question of just *how* the agent should intervene is left open. But, thankfully, the answer is straightforward in principle. When an agent is in a position of being intervention-

sensitive morally responsible, the agent should engage in the most appropriate method of intervention for that entity.

Say you come to find out that you have a racist implicit attitude. The process of identifying the attitude as such is self-directed, oblique blame in which you blame yourself in your relationship to the attitude (you, of all people, are the best person to intervene on your own attitudes!). The appropriate response is then to intervene on such an attitude. To intervene, you look to the best psychological science on implicit bias to see how to shift the attitude or to prevent it from influencing your behavior.

Alternatively, perhaps an organizing group you're involved with finds that a city ordinance in your town is racist. The process of identifying the city ordinance as such is a self-directed, oblique moral evaluation in which you blame yourself for living in a city with such an ordinance. To intervene, you follow the most effective ways for changing city ordinances: writing letters, calling your city council people, showing up to city council meetings, and voting differently in upcoming elections if necessary.

In addition to problems of knowing how best to intervene, it seems that this notion of responsibility introduces of problem of knowing when to intervene. How does S know that E has P? This is an area where moral responsibility and the practices of blame are closely tied together, and another good reason for thinking responsibility should be derived from our accounts of blame. Oblique blame serves as a vehicle for S to know that E has P because it is through blaming T that S shares this knowledge. On its face, it seems that intervention-sensitive moral responsibility has a special knowledge problem, but the same knowledge-sharing process occurs within straightforward blame. When S straightforwardly blames T for φ-ing, S lets T know that T failed to meet a legitimate demand that S had on T when he φ-ed. Just like a friend might say, "You should have told me," when she finds out she was hurting her friend feelings, the primary function of blame is to communicate wrongs and appropriate locations for change. A problem arises, however, because many of the testifiers (S) are People of Color, and as we know from the vast literature on epistemic injustice, the testimonies and resources for understanding developed by People of Color are often excluded as a means of maintaining current power structures. This problem and strategies for shifting these epistemic practices are discussed in Chapter 5.

Our responsibilities to intervene will then be sensitive to the best strategy to intervene on particular entities and best epistemic practices. And, we may find that we cannot go at our

intervention strategies alone in the vast majority of cases, both because the best interventions will require collective action and because identifying when we need to intervene will require knowing more and knowing more will require listening and learning. Thus, perhaps the best way to discharge our responsibilities to intervene is to be actively involved in interracial activist and organizing groups that shift collective epistemic practices (again, see Chapter 5) and collate intervention efforts into a strong base. This account of intervention-sensitive moral responsibility, alongside oblique blame, gives us an understanding of blame that motivates actions. In addition, the account draws attention away from the usual suspects of blame (individuals, their characters, their specific actions) toward features of our social world and the opportunities that individuals have to intervene on the social world toward anti-racist goals.

4.3.6   A Structural Case

Now that I've articulated oblique blame and intervention-sensitive moral responsibility, I'll move into a case that highlights the ways in which oblique blame helps us understanding individuals' responsibilities to intervene into structural injustices. Let's use a slightly revised version of a case from Haslanger's (2015, 6-7, italics my addition) article arguing against individualism:

> Rashaan and Jamal are public high school students in a history class together; race relations in the school and surrounding [sic] are fraught. The teacher, Ms. H., and about three quarters of the class are white. In a discussion of the assigned material, Rashaan repeatedly interrupts Ms. H. to disagree with her and talks over the other students when they try to answer her questions. Ms. H. asks Rashaan to stop interrupting and to wait his turn, but this just makes him more agitated. Eventually Ms. H. asks him to leave and report to the Assistant Principal's office. *After class, Ms. H. overhears Jamal say, "Man, that was racist" to a few of his fellow classmates.* Jamal and other non-white students in the class interpret Ms. H. to be calling out Rashaan because he is Black, and stop trusting her. As a result, they do not engage the material and do poorly in the class.

Haslanger's goal in developing these cases is to show that implicit bias (understood as an individualist mechanism) cannot be the normative core of racism or other social oppressions. She develops the case such that implicit biases don't play a role in the bad outcomes to that structural features of the case, rather than individual features, lead to bad outcomes. For this case, she appeals to the social meanings the students have to interpret Ms. H.'s actions to explain why the students subsequently do more poorly in class. Ultimately, she uses these cases to argue that intervention

into individuals' implicit biases won't lead to the social change at which we're aiming, because implicit bias doesn't play into the causal story in *these* cases. Of course, I'd like to point to the problematic intervention monism here (discussed in Chapter 2). But, more importantly for this chapter, I'd like to use this case to turn more closely to the kinds of moral evaluations that are occurring within this case, especially given the slight modification I've given (one that, at least on my lights, mirrors reality).

In my revised case, Jamal has identified something as racist, but his verbalization doesn't make it clear exactly what entity the predicate applies to. His use of 'that' seems to point to the entire interaction, including, but certainly not limited to, Ms. H.'s unbiased application of the non-disruption policy (Haslanger 2015, 7). If we are to get at Jamal's concerns, we must identify all of the relevant primary entities that are involved in the interaction: the physical and social location of the school, the curriculum, Ms. H.'s non-disruption policy, Ms. H's commitment to the non-disruption policy, Rashaan, Jamal and the other students of color, and the white students.

Let's add into the case that the curriculum misrepresents the history of slavery in the United States. Rashaan, noting this misrepresentation, is speaking over Ms. H. in part because he is frustrated with the curriculum and is unwilling to follow the non-disruption policy due to the mistaken curriculum. Let's assume that in this case the curriculum meets the application conditions laid out in Chapter 3. Thus, when Jamal says, "That's racist!" we can interpret the claim as including within its scope the curriculum. Jamal's statement, on this interpretation, is true. The curriculum is racist.

Now that we've turned up the resolution on the case, we can think more clearly about the processes of blame at play in it. As I argued in the Chapters 2 and 3, the predicate 'racist' can be used to identify that an entity has the property of being racist. However, as Scanlon (1986) argues, when descriptions are about moral features, they have an evaluative charge that is unavoidable. In other words, people feel brought under moral evaluation when it is used, even if it wasn't used to describe them.[81] However, because the curriculum is a non-agential entity—it can't feel the moral pull of an evaluation or do anything to change itself—the moral evaluation and judgements of moral responsibility must occur elsewhere.

---

[81] Of course, some people are exceedingly convinced that they are such good white people that this could never happen to them. And, this is probably more common than this theory suggests. However, like the white man in the bar that feels blamed just overhearing a conversation about racism, I think it is plausible to suggest this occurs very often.

Ms. H. finds herself in a situation where an entity that she has a relationship to—the curriculum—has been evaluated as racist. She herself, as a whole person, has not been evaluated as racist, nor have any of her mental states or motivations (again, if we assume that Jamal is claiming that the curriculum is racist). However, the moral evaluation still seems relevant to her because she is intimately involved with the curriculum, she teaches it, even if she didn't develop it and even if she is required by her school to teach it as it has been developed. Thus, if Ms. H. cares about her students and has some sort of non-racist or anti-racist commitments, she will feel the pull of the moral evaluation even it was directed at the curriculum.

Let's say that Jamal also means to claim that the non-disruption policy is racist and so is Ms. H.'s decision to enforce it (this mirrors another case described by Haslanger (2015) in which a manager fires an employee based on a policy after the employee is unable to get to work on time due to a change in bus schedules). One way to interpret this is to say that Jamal straightforwardly blames Ms. H for failing to meet a legitimate demand he has on her to be treated fairly in light of a racist curriculum. I think this is viable. However, I think a stronger way to interpret the case is that Jamal obliquely blames Ms. H., calling on her to take action and intervene on the policy. One way of doing so would be to not enforce the policy; another would be to try and change it.

When Ms. H. overhears Jamal say, "That's racist" to his friends. She is blamed indirectly (Jamal didn't say it to her), verbally (he used words to express his blame), and obliquely (he evaluated a non-agential entity in which she is in close relationship with but not her). Straightforward blame is insufficient to understand the case. To make sense of Ms. H.'s feelings—even if she reacts negatively, by becoming defensive, for example—and her potential motivation to act, we need the concept of oblique blame. In an ideal case, Ms. H would take this oblique blame and take action toward intervening on the curriculum. Here are a few potential ways of intervening: First, Ms. H. could acknowledge the problems with the curriculum to the class and acknowledge that Rahsaan's resistance is legitimate. Second, she might teach outside of the curriculum and utilize some counternarratives with respect to the legacy of slavery in the United States, even if teaching the curriculum as it was developed is required by her school. Third, perhaps she gathers a few teachers together to review the curriculum and to write a letter to the school district demanding that the curriculum be revised. The possibilities for Ms. H.'s interventions are wide and require interacting with a wide variety of entities.

4.4   Moral Responsibility Debates

The vast literature on moral responsibility in philosophy was spurred by Strawson's ([1962] 1993) famous discussion of free will and the reactive attitudes. As has been clear throughout the paper so far, I have set aside questions of freedom and determinism and have rather utilize some of Strawson's insights, as well as the many revisions and additions, to inform a theory of blame and moral responsibility for racism. Though his goal in the paper is to convince us that the truth value of the determinist thesis has no bearing on the justification of our moral practices, the influence of his *participant reactive attitudes* account is far-reaching and continues to influence accounts of moral responsibility today (Eshleman 2014). In this section, I'll review Strawson's starting point, as well as some key features that inform the Strawsonian moral responsibility literature, in order to show how oblique blame accords and departs from this literature.

Strawson ([1962] 1993)[82] articulates an account of moral responsibility that takes our practices of holding accountable as its starting point. He develops an account of reactive attitudes, or those special human attitudes such as resentment or gratitude, that we direct toward those with whom we are in relationship and more broadly those who *participate* in the moral community. His account is offered primarily to assuage worries about determinism and our moral practices, but this paper has had long lasting effects on the philosophical conversation on moral responsibility beyond this question.

Traditional debates in the free will and moral responsibility literature have assumed that our practices of holding responsible are *justified* in virtue of some fact about whether or not someone is responsible. In other words, we are only justified in praising or blaming due to some fact about the person we wish to praise or blame (e.g. that they could've done otherwise). These debates assume that whatever facts about the world that would justify our moral practices (e.g. the fact that an agent has free will) are outside of, or external to, our practices of holding responsible. Thus, for incompatibilists, determinism raises problems for moral responsibility.

Strawson's goal in articulating his *participant reactive attitudes* account is to show why determinism being true doesn't influence the available justification for our practices of praise and blame. Specifically, he wishes to provide an account that is both unaffected by determinism and that does not rely on a consequentialist view of the justification for our practices of holding each

---

[82] My understanding of Strawson ([1962] 1993) is deeply informed by the OZSW Graduate Course on Moral Psychology that I attended in Delft, Netherlands with Mark Alfano, as well as Eshleman (2014).

other accountable. He ([1962] 1993, 48) writes, "But this [the efficacy of our practices] is not a sufficient basis, it is not even the right *sort* of basis, for these practices as we understand them." Instead, he puts forth an argument in which our practices are justified when an agent deserves to be blamed, but this desert does not come from some independent fact about him (namely his free will).

Rather, Strawson argues that the justification for our practices of holding responsible are within the practices themselves rather than outside the practices (e.g. in some metaphysical fact about free will). He ([1962] 1993, 64) writes that our very human practice "neither calls for nor permits, an external 'rational' justification." This is because our practices arise out of desires to be treated in particular ways and desires are neither true or false. He thinks this shows that determinism's truth is irrelevant to our practices of holding responsible. Thus, in our thinking about moral responsibility, we should focus on our "ordinary practices" of holding responsible rather than on these deep metaphysical questions about freedom (Knobe and Doris 2012).

In the rest of this section, I'll focus on three topics within the Strawsonian blame literature: ordinary practice, the ethics of blame, and attributability. My account of oblique blame takes a position with respect to these debates and some clarification is necessary. It is my goal in this section to answer some of the natural questions that arise from this account and to situate oblique blame with respect to these longstanding lessons and debates.

4.4.1 "Ordinary Practice"

One of Strawson's ([1962] 1993) primary goals is to shift the conversation around moral responsibility from questions of free will and determinism to looking at the ordinary practice of moral appraisal, specifically our practices of praise and blame. It is from these practices that we should be able to learn something about moral responsibility. Though I'd like my account to be sensitive to ordinary practices of holding accountable, my project in this chapter is revisionist. In other words, I think there is something wrong our current practices of holding each other responsible with respect to racism. So, though my theory isn't committed to unification or invariantism—there may be many kinds of moral responsibility and there could be a wide variety of criteria we use in different situations for making judgments about moral responsibility—it may not comport with people's ordinary judgements.

The reason for my revisionist account here is twofold: first, ordinary practice differs between groups of people; what is present in ordinary practice for some may not be for others. I take this to be a general worry about making universal claims about moral responsibility from ordinary practice, even when it is informed by empirical studies, since the ordinary practices studied are often from a particular social location. But my caution here is also informed by two empirical studies that found race-based differences in moral judgements about racism. Cameron, Payne, and Knobe (2010) found [83] that African American participants attributed moral responsibility to individuals for implicit racial biases more often than Caucasian[84] participants, even when the biases were described as unconscious. In a review of studies comparing race-based differences in perceptions of racism, Carter and Murphy (2015) found that there are differences between white and Black peoples' intuitions about the attitudes and behaviors that count as racism.

The second reason why we need a revisionist account is epistemological. In his seminal book, Mills (1997) argues that white people have adopted an epistemology of ignorance for matters related to race. This epistemological contract requires that whites misinterpret the world such that their ordinary judgements are flawed. He (1997, 93) writes:

> Whites will then act in racist ways while thinking of themselves as acting morally. In other words, they will experience genuine cognitive difficulties in recognizing certain patterns *as* racist, so that quite apart from questions of motivation and bad faith they will be morally handicapped simply from the conceptual point of view in seeing and doing the right thing. As I emphasized at the start, the Racial Contract prescribes, as a condition for membership in the polity, an epistemology of ignorance.

Due to the epistemology of ignorance, looking at ordinary practices of white practices of moral evaluation and blame won't tell us much about how we should go about holding each other accountable. Seen through a revisionist lens, "ordinary practices" of blame are simply one type of blame that magnifies individual confrontation and resolution. Taking a revisionist approach means not only broadening from that focus, but also understanding that straightforward blame makes responsibility for structural injustices difficult. In other words, this revisionist lens shows that including more than our ordinary practices of blame is necessary when these ordinary practices

---

[83] Cameron, Payne, and Knobe caution that these results should be interpreted with caution since the sample size of African Americans was very low.

[84] In my descriptions of both studies, I'm using the racialized category names used in the studies.

are meant to obscure responsibility for racism (for more on anti-racist epistemological work as it relates to interventions, see Chapter 5).

Thus, rather than looking only to ordinary practice, I aim to offer a new way of thinking about moral responsibility for racism, a new set of criteria for moral judgments with respect to racism, so we can do a better job of holding each other accountable (and a better job of allowing ourselves to be held accountable) and shaping the world in a way that reduces the number of racists things in the world and reduces needless suffering and death. In this away, my account of oblique blame and intervention-sensitive moral responsibility is a normative addition to the literature. It doesn't conform to current practices (and it certainly changes what we think it means to be blamed), but rather it suggests we change our current practices such that oblique blame is a viable practice of accountability. This changes the standards on which my account should be evaluated from how we *do* blame to how we might *improve* blame and responsibility.

4.4.2   The Ethics of Blame

Those who have taken up the distinction between moral responsibility, blameworthiness, and blame (discussed above in my description of straightforward blame), have raised another question: When is appropriate to blame? In other words, when would blame be deserved, good, permissible, or fair (Coates and Tognazzini 2013)? In addition, these theorists think that there are times when the agent is blameworthy (deserving of blame), but it would still be inappropriate to blame them. For example, it would be unfair for a liar to blame another liar for lying, and it might not be good to blame someone that is deserving but on the brink of a mental breakdown. These same considerations apply to oblique blame, but an additional concern arises. Namely, when is blame too demanding? When is one's intervention-sensitive responsibility discharged?

My account of intervention-sensitive moral responsibility says that T should intervene on an entity (E) when (1) E has a negative moral property (P) and (2) T is related to E in a way such that T's intervention may be effective. In a world such as ours, where a great many things will be racist, this leaves individuals in a position where they are very often responsible in this intervention-sensitive sense. This leaves the individual open to oblique blame in any case where another person (S) judges that T is related to E (but not identical to E) such that T's intervention on E may be effective. Given the ubiquity of individuals' intervention-sensitive moral

responsibilities and the deserved blame to which they will be open given my theory, we may think that this theory of moral responsibility is too demanding.

Included in the conditions for blame is that T's intervention on E may be effective. Effectiveness here captures a wide variety of things, including both the resilience of E and the qualities of T. Since the resilience of E is somewhat independent from our practices of blame, I'll focus on the qualities of T. These qualities include the knowledge of T, the skills of T, the interests of T, and T's relative social position in a given society. For example, let's imagine a target of blame who is a dreadfully poor public speaker and an opportunity to intervene that would be effective given a wonderful public speaker. In this case, T would be an inappropriate person to blame since he will be ineffective in intervening on E. Perhaps most important is T's social position in the current structure of society, even at a local level.

Let's go back to the classroom example. Say that Rahsaan continued to speak over Ms. H because he had directed oblique blame at himself: he was trying to exercise his responsibility to intervene on the racist curriculum. However, because of his position relative to other actors, his intervention wasn't very effective. Ms. H. then is a more appropriate target, and even more appropriate than Ms. H. are school administrators and other people in a position to change policies. Note, however, that intervention-sensitive moral responsibility and oblique blame can have a cascading effect. Even if Ms. H. isn't the *most* effective target of oblique blame, her reaction to the blame causes her to hold other people to account by directing oblique blame at them. These people may do the same until oblique blame reaches the most effective target. Thus, oblique blame can still be appropriate even in cases were T isn't in an optimal position to enact change.

When we consider whether it is appropriate to obliquely blame some, then, we'll think about the qualities they possess and take these into consideration when we consider a given target of blame's effectiveness. In addition, we may also take into consideration the values of time and balance with respect to demandingness. Like all of us, a potential target of blame will have a great many responsibilities, including responsibilities to his family and friends, professional responsibilities, and other responsibilities to herself. Thus, in our practices of blame, we should keep these in mind and blame in line with time and balance with respect to a target's other moral responsibilities and needs.

4.4.3 Attributability

From Aristotle ([384-323 BCE] 1984, *NE* 3.1-5) to very many contemporary accounts, an agent must be related to an action in some appropriate way for praise or blame to be appropriate. For Aristotle, for an action to be worthy of praise or blame, the action had to be *voluntary*. Contemporary theorists cash this out in terms of control and attributability. Those committed to control maintain that for an agent to be responsible for an action, then she must have been in control over it (e.g. Vargas 2013). Other views, which want to leave open the prospect of responsibility for non-voluntary actions, have developed attributability views. These views follow from a tradition of "deep self" views (Wolf 1990) and include an account of how actions "belong" to an agents. As discussed above, on Smith's (2007) view, an action is attributable to an agent when that action has a rational connection to an individual's evaluative judgements.

This aspect of moral responsibility has been used to argue that agents are not morally responsible for things like implicit bias (Levy 2016, 2017) and causes immediate problems for things like structural injustice since these injustices cannot be attributed to one agent. We see this basic argument when it comes to reparations for genocide, slavery, and internment: *I* didn't do any of those things; why should *I* be held responsible?

However, there have been some efforts to assuage these worries. For example, Brownstein (2016, 2018) has argued that implicit biases are attributable to agents in that they are reflective of our cares. Others have argued that though we aren't directly responsibility for things like implicit bias, we are indirectly responsible for doing things to shift biases or limit their influence on our actions (Holroyd 2012), especially when we should've known that implicit biases influence actions (Washington and Kelly 2016).

Other approaches avoid problems of attributability by arguing that we can still be held accountable without being blameworthy or open to being blamed. For example, Zheng (2016) argues that implicit biases aren't attributable to agents (unless they are endorsed), but that people can still be held responsible for them by being held accountable for doing something to address the bad effects (but not blamed). Zheng (2018) takes a similar strategy when she argues that individuals should be held accountable for structural injustice. Similarly, Mason (2019) argues that we can extend responsibility to ourselves by *taking* responsibility for things that we did not intend, but have bad effects. Our motivation for doing so are the relationships that may be affected by the effects of our actions. Though these strategies also avoid problems of attributability, they also

sacrifice blame along with it. This is a problem because, as noted earlier, blame plays an important knowledge-producing role in responsibility practices.

My account of oblique blame side steps questions of attributability while maintaining the motivating function of blame, as discussed in the *desiderata* for an account of moral responsibility for racism. Though this feature is what makes my account of moral responsibility and blame so very different from straightforward blame, it is also what helps us understand blame for a whole host of inherited social injustices and problems that individuals in this lifetime didn't cause and do not knowingly maintain (because of the epistemology of ignorance and forms of willful ignorance (Pohlhaus 2012)). Oblique blame highlights the ways in which attributability is a distraction in certain cases. It's arguing over who caused a disaster when the most important thing to do is to get out of it. Thus, while Levy (2017) would count my disinterest in attributability against me, oblique blame recharacterizes the attributability debate as overly focused on maintaining the innocence of individuals at the loss of taking responsibility together for robust change (see also McHugh and Davidson forthcoming).

4.5   Conclusion

Oblique blame has the power to transform our practices of holding each other accountable for racism. As confusion around the predicate increases and individuals hold onto the desire to be "good people" ever so tightly, the concept of oblique blame provides a pathway forward. In staying true to entity type pluralism, we are able to see an expansion of the types of entities that can be described as racist without losing our ability to take action with respect to racism. As I mention in Chapter 2, structuralists care about structures because they think structures play the biggest role in the maintenance and development of racist outcomes. However, when the normative core of racism is located in these non-agential structures, the pathway for individuals can become unclear. My account of oblique blame provides a clear path for individual accountability in the midst of complicated social structures.

However, oblique blame only works insofar as we are able to identify the entities that have the property *racist*. Even with the application conditions in hand, we may have a hard time figuring out whether not a thing is racist. However, there are some people, people who experience racialized violence, who are in a better epistemic position to identify racist entities (Mills 1998), especially those entities that are causing them and their friends and families harm. As I've argued elsewhere

(Davidson manuscript), these testimonies are not always given the credibility they deserve due to contributory epistemic injustices (Dotson 2012) that track racist epistemic resources. Thus, a closed loop is introduced: those who are in the best epistemic position to identify racist entities aren't believed because of racist entities like racist epistemic resources.

In Chapter 5, I develop a strategy for combatting this closed loop. The development and maintenance of epistemic resources is governed by epistemic norms and social practices, which are both governed by the mechanisms of social norms. When epistemic resources are inadequate (read: racist), one way to intervene on these resources is by changing the norms that govern their production. I utilize the literature on social norms to frame this conversation and to develop an account of social norm intervention for epistemic resources. My development of this chapter is an attempt to direct oblique blame at myself to motivate me to identify strategies for intervening on the epistemic resources that govern our ways of thinking, and thus our ways of being.

CHAPTER 5.   LIBERATORY (EPISTEMIC NORMS): ADRESSING THE KNOWLEDGE PROBLEM FOR INTERVENTION-SENSITIVE MORAL RESPONSIBILITY

That popish recusants convict, negroes, mulattoes and Indian servants, and others, not being christians, shall be deemed and taken to be persons incapable in law, to be witnesses in any cases whatsoever. – Laws of Virginia, Chap. XIX.XXXI, October 1705

Section 394 of the Civil Practice Act provides, "No Indian or Negro shall be allowed to testify as a witness in any action in which a White person is a party." Held, that the words, Indian, Negro, Black and White, are generic terms, designating race. That, therefore, Chinese and all other people not white are included in the prohibition from being witnesses against Whites. - People v. Hall, 4 Cal 399 (1854)

5.1   Unreliable Testifiers

From coast to coast in the United States, there were laws in the eighteenth and nineteenth centuries prohibiting the testimony of People of Color from being used in courts when a white person was involved with the case. This included cases in which the white person was the perpetrator of violence and murder and People of Color were the only witnesses to the crimes. According to a (1864, 14) Senate Report, the cited justification for these laws was the "the degraded condition of the slave, and the interest which he may have to conceal or deny the truth (Wheeler's Laws of Slavery)." The report (1864, 15) goes on to claim, however, that the laws which prohibit testimony are not limited to enslaved persons (nor those who are not religious), and thus, the inability for People of Color to testify is merely "an incapacity attached by law to persons of color." In other words, the report argues that all attempts to justify such laws are inconsistent with the practice of such laws, meaning that the exclusion of testimony is based on racialization alone.

The laws that exclude People of Color from giving testimony in court in cases involving white people have long since been repealed. However, that doesn't mean that the practices that brought about the codification of the practices into law do not persist or that these laws don't have long lasting effects on epistemic norms (for an example, see Acharya, Blackwell, and Sen (2018)).

For an example also within the court room, *Batson v. Kentucky (1986)* outlaws the state from using its peremptory challenges to dismiss jurors based on race, but it is still common practice to dismiss jurors using cited reasons other than race while still targeting potential jurors of color. In fact, the DA has provided a "cheat sheet" to its prosecutors to get around the *Batson* ruling.[85] Some of the reasons listed are relevant to epistemic matters, e.g. attitude, body language, juror responses, communication difficulties, and signs of defiance or antagonism toward the State, while others are purely coded-racial language, e.g. physical appearance. The continued exclusion of jurors based on race is clearly tied to the specific outcomes desired by the parties involved, but in addition it tells us that our collective epistemic practices often exclude the testimonies of People of Color.

Though these practices of exclusion are easily seen in the history of the legal system because of the systematic and public ways these exclusions have been recorded, these exclusions are present throughout dominant epistemic practices. Lowered credibility based on the racialized identify of the speaker alone are race-based identity-prejudicial credibility deficits (Fricker 2007, 28), a specific type of epistemic injustice in which a speaker is wronged in their capacity as a knower because their testimony is given less credibility than it deserves or when their status as knower is denied all together. When a persistent exclusion from the knowledge production community occurs, the result is epistemic oppression (Dotson 2014). These exclusions are prescribed[86] by our current epistemic norms, which stabilize our epistemic practices. Our current epistemic practices lead to the epistemic oppression of People of Color.

My discussion will proceed on the uncontroversial premise that epistemic oppression is wrong in itself, which gives us reasons to intervene on epistemic practices that lead to and uphold these oppressions. However, in this chapter, I'm interested specifically in the ways patterns in our epistemic practices raise problems in our collective ability to identify instances of racism. People of Color, those with racialized identities who experience racism, will be in the best epistemic position to identify when racism is showing up in our dominant ways of thinking and in our interactions, norms, and formal institutions. This is reflective of the thesis of epistemic advantage (Jagger 1988; Harding 1991; Mills 1998; Wylie 2004; Rolin 2009). Those who endorse this thesis argue that individuals have unique knowledges that arise from their social experiences, and in

---

[85] https://www.aclu.org/legal-document/north-carolina-v-tilmon-golphin-christina-walters-and-quintel-augustine-batson

[86] I use the word 'prescribed' here to align myself with Mills (1997). He argues that the Racial Contract prescribes certain patterns of unknowing, certain epistemic practices, that cover over matters related to race.

particular their experiences of oppression. Jagger (1988, 370) writes, "Their [oppressed people's] pain provides them with a motivation for finding out what is wrong, for criticizing accepted interpretations of reality and for developing new and less distorted ways of understanding the world." It's important to note that epistemic advantage does not require that all people in a particular position *will* know. It is not a matter of biological causation or social determination, but rather of being in a position to know given certain similarly structured social experiences (Mills 1998). Alternatively, those who are in this position of epistemic advantage may adopt an adaptive preference (Khader 2011), a preference that arises from unjust social conditions, to not know because knowing more in these cases is difficult and painful.

If we assume the thesis of epistemic advantage, that those who have experienced racism regularly are those who are in the best position to know when it is occurring, then the exclusion of People of Color from the knowledge production community means that our collective epistemic resources are both inadequate (Dotson 2014) and unjust. When the testimonies of People of Color are dismissed or not taken seriously due to epistemic oppressions, important information about the experience and nature of racism is excluded from the dominant epistemic resources or ways of knowing. Further, this exclusion means that the testimony of People of Color is not always intelligible within the dominant epistemic framework, meaning it is more likely the testimony will be excluded. As Medina (2012) argues, the gaps in our collective epistemic resources are due to testimonial injustices, but testimonial injustices cannot be addressed without expanding our collective resources for understanding. Testimony and epistemic resources are in a tight feedback loop, making interventions into the norms that govern our epistemic practices difficult.

Throughout this chapter, I'll discuss "our" epistemic practices. Because epistemic norms and practices are localized to particular populations, in a general sense, I'll be using "our" to talk about epistemic practices within United States, since that's the context with which I'm most familiar. But, in addition, when I use "our," I'll be talking about dominant epistemic practices, those practices that are common, utilized, and maintained by those who have the most social and political power. My critiques are not directed toward subaltern epistemic practices that have been developed within communities that the dominantly situated oppress both epistemically and materially. There are clearly very valuable epistemic practice shifts and resources developed within these non-dominant and epistemically oppressed groups (as well be discussed in Section 5.5). However, due to the pervasive and invasive nature of dominant practices, even those who have developed

resources for understanding outside of the oppressive dominant practices will be subject to them (Lugones 1987). Thus, in an important sense, the dominant resources are "our" resources. This line of thinking is informed by Pohlhaus' (2012) articulation of *willful hermeneutical ignorance* in which the dominantly situated willfully refuse to take up the epistemic resources developed by those they have oppressed in order to maintain current epistemic and material systems that benefit them. The resources developed in this chapter are meant to begin to push against willful hermeneutical ignorance[87] and deconstruct the ignorance inherent to our dominant epistemic resources through intervention into epistemic norms.

In the previous chapters of this dissertation, I have argued that increasing the conceptual clarity around racism, especially the predicate 'racist', will allow us to preserve the moral function of the predicate and better hold each other to account for racism In Chapter 4, I put forth an intervention-sensitive theory of responsibility that allows us to understand the ways in which moral agents can be responsible for intervening on non-agential racist entities. I ended that chapter by raising a problem: intervention-sensitive responsibility only works if entities can be correctly identified as racist, and given that our epistemic norms and resources already serve to cover over racism (Mills 1997), we will continue to have epistemic problems even once the concept of racism has been clarified. Thus, there are serious epistemic barriers for intervention-sensitive responsibility. Throughout this chapter, I'll call this the Knowledge Problem for Intervention-Sensitive Moral Responsibility, or the Knowledge Problem, for short.

If we want to take the Knowledge Problem seriously, we're going to have to look closely at our epistemic norms and practices. Thus, in this chapter, I argue that one promising arm of the anti-racist project focuses on shifting the norms that govern knowledge production and the building and changing of epistemic resources. Because the account of oblique blame and intervention-sensitive moral responsibility I've articulated is focused on effective intervention, I'll utilize the empirically-informed philosophical research on norms and norm change to frame the intervention into these epistemic practices. In other words, this chapter has both theoretical and practical aims.

---

[87] An interesting tension here is how theories of norm psychology will square with the idea of willful ignorance. What motivates Pohlhaus (2012) to describe this ignorance as actively constructed is her desire to hold individuals responsible (likely because of the attributability conditions on traditional account of moral responsibility, see Chapter 4 for more discussion). However, as I'll explore further in the next section, I tend to think that norms can override the will in some cases. In other words, even if I have a commitment to deconstructing ignorance, I may engage in practices that maintain it due to the norms that govern our collective epistemic practices. This is the complex issue with which this chapter is concerned.

On the one hand, I want to offer a theoretical account of how epistemic norms are related to the Knowledge Problem and the maintenance of racism. And on the other, I want to offer practical tools for intervening on norms that maintain epistemic systems.

Here's how the paper is organized. In Section 5.2, I'll outline the contemporary research on social norms, focusing on Bicchieri's social expectation view (2006, 2017), as well as the cognitive evolutionary view of norms presented by Kelly and Davis (2018) (see also Sripada and Stich 2007; Richerson and Boyd 2008; Henrich 2016). In Section 5.3, I situate my project with respect to the current work on epistemic norms within mainstream analytic epistemology. In Section 5.4, I'll connect norms to epistemic practices, showing some ways in which epistemic practices are governed by norms, and show the ways in which dominant practices serve to obscure thinking about racism contributing to the Knowledge Problem. Finally, in Section 5.5, I'll discuss four candidate norms, as well as the formal structures that support their incubation, developed within activist and organizing spaces that are designed to be interventions into epistemic practices and to disrupt dominant knowledge systems. These norms are liberatory norms in their capacity to dismantle epistemic systems that support racism. These candidate norms focus on the process over the product, centering those most affected by an issue, progressive stack, and *The Circle Way* (a facilitation process that structures conversations). I'll conclude with some thoughts about how these candidate norms might be installed into larger communities. This chapter connects the literature on epistemic oppression to work on norms and norms intervention to provide a concrete pathway for addressing the Knowledge Problem for moral responsibility.

5.2   Norms

Norms make up a family of social mechanisms that are often appealed to in order to explain behavior. Norms govern a whole host of our social interactions from how we interact with other cars when a two-lane highway goes down to one-lane for construction to the way a silent hush falls over the theatre when the lights go down to the way that men can catcall women in the streets without any negative response from their friends. On its thinnest conception, norms are patterns of behavior, what an individual person in a certain situation is likely to be doing given the collective-level patterns of behavior in the group in which that individual finds themselves. However, philosophers (and social scientists) have sought to provide richer conceptions, conceptions that can capture the ways in which we are pulled to behave in accordance with norms and are compelled

to sanction others when they don't conform to norms. Some of these aspire to explain how they can do so regardless of individual's avowed beliefs, and the reasons why patterns of behavior can be stabilized over time even when they are maladaptive or unpleasant. This latter sense of norms is what I'll be focused on in this chapter.

5.2.1 Bicchieri's Social Expectations View

In her (2017) book, *Norms in the Wild*, Bicchieri lays out a such an account, focusing on what she calls *social norms*. For Bicchieri (2017, 35), a social norm is a:

> rule of behavior such that individuals prefer to conform to it on condition that they believe that (a) most people in their reference network conform to it (empirical expectation), and (b) that most people in their reference network believe they ought to conform to it (normative expectation).

What's distinctive about this view is that what makes a group-level pattern of behavior a social norm is that it is stabilized by a particular set of psychological states had by individuals within the group. This is important because interventions into social norms will be focused on changing these specific psychological states, focused on changing people's empirical expectations—expectations about how others *will* behave—or their normative expectations—expectations about how others think people *should* behave. Empirical and normative expectations support *conditional preferences* or preferences for behaving in a particular way given that a person thinks others will behave that way (empirical expectation) or thinks others think people should behave that way (normative expectation). In addition, norms are localized to *reference networks* or the groups of people that we care about when we are making decisions about what we should do.

In this book, Bicchieri is narrowly focused on *social* norms, or those norms that are stabilized by both empirical and normative expectations. Bicchieri distinguishes customs, descriptive norms, and moral norms from her account of social norms (the primary focus of her book). However, because I'm interested in norms in general here, I'll review each one in term. By her lights, customs are those behavioral patterns which are sustained by individual's independent motivations stemming from their prudential needs. For example, it is customary to use an umbrella when it is raining, but I don't use one because I think others will and because I think others think I should, but because I don't want to get wet. However, Bicchieri thinks that when it comes to customs with negative effects, social norms may be integral to changing these customs, since social norms have the power to motivate individuals from doing what is habitual or easy.

What Bicchieri calls descriptive norms, on the other hand, involve conditional preferences to act based on empirical expectations alone. Thus, one conforms to a descriptive norm because they believe that most people in their reference network conform to the norm. However, descriptive norms lack the normative component. Thus, changing descriptive norms requires changing beliefs about what other people will do. This can be accomplished by either correcting false beliefs about what others will do or by changing the behaviors of a few influential people in the reference network. Since descriptive norms are not supported by normative expectations, beliefs about what other people think you should do, descriptive norm change doesn't require a change in normative expectations. Finally, Bicchieri asserts that there is an important difference between social norms and moral norms in that the motivation to follow a moral norm stems from one's personal moral commitment, rather than an empirical or normative expectation.

I've started with Bicchieri's social expectations account here because it gives us some straight-forward vocabulary to use when talking about norms on the individual- and group-levels. Though Bicchieri is laser-focused on social norms as the most interesting and important for intervention, these other kinds of norms help us get a sense of the options and the ways in which they many be differently supported by psychological features. Another account, one focused on norms more broadly, will help us think about the larger class of norms.

5.2.2 Kelly and Davis' Cognitive Evolutionary Approach to Norms

In order to focus more expansively on the features that tie norms together and make them a powerful explanatory mechanism, I'll discuss Kelly and Davis' (2018) *cognitive evolutionary* approach to norms. This broader account (2018, 58) is designed to accommodate "norms of logic, language, epistemology, aesthetics, religion, or etiquette" and allows us to see what norms have in common and think more about the motivational and stabilization roles norms play in our cooperative societies. Here are the most useful insights of this broader cognitive evolutionary account:

First, individual humans have a psychological norm system. The system is a genetically and culturally evolved set of mechanisms that allow humans to pick out norms in our local environments, remember and internalize them, and sanction people (including themselves) when the norms aren't followed. The capacities within the norm system are cross-culturally present, but the content of norms (what we ought to do and how we ought to sanction) varies widely from

culture to culture. Stemming from Sripada and Stich's (2007) seminal two-level framework for understanding norms, many accounts of norms systems include both a mental or representational component and a behavioral component (Kelly and Davis 2018, 61), i.e. an account of how norms are "stored" within individual minds and an account of how this influences behaviors on both the individual and collective levels. In addition, the norm system operates outside of the realm of conscious or deliberative control, by System-1 processes.[88]

Second, norms have a "normative force" (Kelly and Davis 2018, 61). This means that once a norm has been picked up by an individual's norm system, they will be pulled to enforce the norm on themselves and other people even if they do not avow the norm. This enforcement consists in social sanction or reward. For example, I sometimes find myself reacting negatively to woman's leg hair (a form of social sanction), even though I do not remove my own leg hair. The nature of the norm system makes norms that have been picked up, or internalized, by an individual feel natural; they often even become integrated into the inner voice that guides behavior. The internalization of the normative force means that individuals follow norms not just because they can avoid sanction by doing so, but because they feel it is the right thing to do. This does not mean that an individual will always act in accordance with a norm, rather it means that they have an intrinsic motivation to comply with the norm and punish others who do not, but the intrinsic motivation to comply could be outcompeted by other motivations for an individual. The normative force outlined by Kelly and Davis is broader than the social expectations account given by Bicchieri in that it captures the felt sense of justice that isn't captured by the normative expectations condition in her social norms account.

These two features help us understand how norms function in general by specifying that humans have a special capacity to absorb, to socially learn, the norms in their particular environment and by articulating what it means for all norms to have a normative force, a pull to conform and reward or punish others who comply or rebel. These two mechanisms help explain why norms are stabilized and why changing norms, especially on purpose, can be difficult. Another important feature for my purposes is that this broader view of the function of norms captures features of epistemic norms, which may or may not be captured by Biccheri's view, depending on the psychological content of such norms. I've reviewed the norm literature here because it is deeply

---

[88] For an overview of these quick, pre-reflective, and automatic processes and how they differ from deliberative, conscious reasoning, see Kahneman (2013).

rooted in empirical studies of how norms function and the ways in which they might be changed. This is important since effective intervention is a key feature of intervention-sensitive moral responsibility and addressing the Knowledge Problem will require intervention into epistemic norms. This is the first step in knitting together the theoretical account given in Chapter 4 with the practical and concrete account I give in this chapter.

5.3 "Epistemic Norms"

In the previous section, I gave an overview of norms, the processes of their stabilization, and the motivation to comply and sanction. This was a broad overview of all norms, but this chapter is about epistemic norms in particular. In recent years, there's been an increased focus in epistemology on epistemic normativity. For example, there's been a growing consensus that our epistemic interactions are themselves governed by norms of assertion: we should only assert what we know or have good reasons to believe is true. The debates within this literature are primarily focused on questions like whether assertion and epistemic actions (like practical reasoning) are governed by the same or two unique norms (Brown 2012; Montminy 2013; McKinnon 2012; Gerken 2014; Kauppinen 2018) and whether friendship requires that we shirk certain epistemic norms (Keller 2004; Stroud 2006; Kawall 2013). In addition, the literature on epistemic norms makes a distinction between truly epistemic norms and moral norms that govern epistemic matters. For example, Kauppinen (2018) points out that norms against lying are certainly about epistemic matters, but that the norm is moral rather than epistemic.

There's an interesting set of puzzles here, and what I'm focusing on in this paper isn't completely divorced from these concerns; however, I'll be interested in the broad category of norms that govern our epistemic practices, which will go far beyond the norm of assertion or norms that govern practical reasoning. My reasons for this are twofold. First, it is a reflection of my commitment to non-ideal theory in general (Mills 2005) and non-ideal epistemology in particular. According to Barker, Crerar, and Goetze (2018), "This kind of epistemology focuses not on what our epistemic lives look like when everything runs as it should—on the nature of justification, the sources of knowledge, or the mechanisms of testimony and trust—but on what our epistemic lives look like when things go wrong" (2). The discussions of epistemic normativity I discussed above are concerned primarily with ideal epistemology, so it shouldn't be surprising

that my discussion here will be far broader and less concerned with finding, for example, the one norm that governs all epistemic matters.

Second, I'll reject the attempt to divide norms into pure epistemic and moral categories and focus on the broader set of norms that govern epistemic practices. Kauppinen (2018), for example, maintains a distinction between moral and epistemic virtue such that one can be morally virtuous without being epistemically virtuous and vice versa. Though we can think of hypothetical cases were this distinction holds, the primary cases with which I'll be concerned in this chapter have to do with the intersection of moral and epistemic concerns. This is first a reflection of my deep skepticism about the separability of moral and epistemic matters in general (Davidson 2017), but also of the deep disagreements around exactly what delineates moral norms from other norms. For all the work that has been done to separate norms into categories, distinctively moral norms have caused philosophers a lot of trouble (Davis and Kelly 2018; Kelly forthcoming). So, in this chapter, I won't rely on such a distinction. Rather, I'll talk about 'epistemic norms' as those norms that govern a wide range of epistemic practices, and though I think this constitutes a family of norms, I'm not committed to these epistemic norms as a natural kind of norms.

In many ways, my broad focus on norms here, especially as it relates to epistemic normativity, will depart quite sharply from the issues highlighted in mainstream epistemology. This is in part due to my philosophical commitment to non-ideal epistemology and the rejection of the distinction between epistemic and moral norms. But, this is also because I'd like my discussion to be attuned to the social and psychological mechanisms that stabilize norms discussed in the previous section, something left out of the philosophical literature on epistemic normativity within mainstream epistemology.

5.4   Norms and Epistemic Practices

In this section, I will dive a little deeper into why norms matter for epistemic practices. Epistemic practices[89] are all those individual and group-level capacities, behaviors, ways of being and knowing, and social structures that make up knowledge production. Epistemic practices can

---

[89] As noted above, the focus on epistemic practices is reflective of my commitment to non-ideal epistemology, which is informed by feminist, standpoint, and decolonial epistemology. The contributions of these projects are vast, but the one I take to be central to this chapter is the shift from focusing on what an individual can know independently of their context to aspects of epistemic practices and the ways in which these practices lead to un-knowing, hermeneutical gaps, and silencing (Mills 1997; Pohlhaus 2012; Dotson 2011).

be very local, e.g. the publication process for a particular journal, or they can be broader, e.g. the inverted epistemology mandated by the racial contract (Mills 1997). Epistemic practices encompass a large swath of very different features and processes, and yet, norms govern and stabilize many of these practices.[90]

But, when we think of behaviors stabilized by norms, we often think about concrete patterns of behavior: how we interact with each other in the line at the grocery store, how we behave in meetings and on the subway, how we drive our cars, and how we interact with friends, potential love interests, and long-term partners. However, norms do also govern and stabilize our epistemic practices, which are made up of both concrete behaviors (listening with attention to the person at the podium) and more abstract or subtle behaviors (pre-reflectively assigning credibility to a speaker). And, it's good that norms govern these behaviors. It's good that there's a norm against murdering people (and good that there's a formal institution[91] that also serves to support this norm). And its also good that there's a norm that says we should only assert things we know or have good reasons to believe are true. Norms govern our epistemic behaviors just as they do other patterns of behavior.

Though norms have a normative force—they pull us to comply with them and to force compliance on others through social sanction—this doesn't mean that norms do not violate standards of justice. Norms can uphold oppressive structures. Further, just because one recognizes that upholding a particular norm violates another's right, for example, doesn't mean that one will violate the norm. In discussing sexist norms, Witt (2011) notes that woman will feel the pull of oppressive norms that they strongly disagree with and will still certainly be punished by others for violating such norms (46-47). Thus, just like we should analyze other behaviors governed by norms, we should analyze our epistemic norms in order to see if and which norms violate standards of justice and block our path toward a more just and equitable world. For the purposes of this chapter, I'm interested in the unjust norms that maintain the Knowledge Problem, those norms that

---

[90] There can also be a meta-stabilization of a formalized epistemic practice. Say, for example, that one believes that the paper review for a particular conference results in mostly white men being accepted even when these aren't the best papers. The review process seems to be the culprit. But, imagine that as one begins to write an email to point out this pattern, one feels the pull to "not rock the boat" or to "not accuse anyone of being racist or sexist." This is the pull of a norm, a norm that mandates compliance with current epistemic practices.

[91] For a discussion of informal institutions (norms) and formal institutions (policies and laws), see (Davidson and Kelly 2018). As discussed in footnote 4, informal institutions can stabilize formal institutions by keeping individuals from protesting them.

govern our epistemic systems that lead to gaps in our epistemic resources such that we are unable to identify instances of racism within our dominant epistemic systems.

Norms stabilize epistemic systems and can have cyclical effects. Medina (2012) emphasizes this when he discusses the tight relationship between testimonial and hermeneutical injustices. Women, for example, are often expected to be quiet and to not express their opinions or ideas too assertively. When this norm is followed or enforced, then we collectively miss out on what women have to say. Thus, this epistemic practice, stabilized by norms about how women should behave, persistently excludes women from or severely limits their participation in knowledge production (Dotson 2014). This means what we know and how we think is missing their perspectives. Further, because of this exclusion, we then have a hard time interpreting women if/when they do speak due to these missing epistemic resources. This reinforces the belief that women shouldn't speak up because they are unintelligible (due to a gap in our epistemic resources).

The norms that govern epistemic practices are tied to other social oppressions. DiAngelo (2018) is a white anti-racist scholar who holds workshops for white people to begin to unlearn their racist ways of thinking, behaving, and interacting. Through this work, she has identified the unspoken "rules of engagement," which are norms that white people enforce during conversations about race. These "rules" serve to protect white ways of thinking and understanding the world such that racism cannot be analyzed or made intelligible. The first rule is "Do not give me feedback on my racism under any circumstances." This rule is a norm that governs epistemic practices in conversations about race and racism that silences People of Color by sanctioning them for giving their testimony. Collectively, we miss out on the knowledge that would be provided through testimony and white people continue to believe that they are "not racist" because they get limited evidence to the contrary. Sanctions for violation of this norm range from social ostracism to severe backlash and violence.

Norms and stereotypes work together to keep testimony unintelligible or inadmissible. When the testimony of People of Color is persistently excluded from epistemic practices, the resources available for understanding are insufficient for understanding these testimonies. In addition, our ideas about reliable knowers also contributes to this problem. The norm of "rationality" or the idea that reliable testimony will be offered without emotion, alongside the stereotypes of "angry Black people" or "feisty Latinxs," serves to silence People of Color by rendering their testimonies unintelligible (can't be understood given current epistemic resources)

or inadmissible (not seen as evidence at all). The norms around credible testifiers and the norm that People of Color should not give white people feedback on their racism govern our epistemic practices such that our collective epistemic resources are inadequate. This both creates and maintains the Knowledge Problem in that instances of racism are continuously covered over, and our intervention-sensitive responsibilities are left unidentified and further unenacted.

5.5 Local Norm Change, Trendsetters, and Norm Installation

In the previous section we saw how norms govern our epistemic practices. Norms can work to stabilize inadequate and unjust epistemic systems[92] such that our epistemic resources are lacking and there is no clear way to develop new epistemic resources because the system is self-preserving. When our epistemic practices lead to insufficient epistemic systems, or systems that lack the resources for understanding the experience of everyone within the system and contribute to the Knowledge Problem, norms can be a location for intervention in the anti-racist project. Though I don't take this to be a controversial point, norms as a location for intervention is often missed within the individualist versus structuralist debate discussed in Chapter 2. For a recent discussion of the ways in which norms have the power to illuminate mediating features that uses ideas drawn from much of the same literature I'm using here—features that are both individual and structural— see Davidson and Kelly (2018).

After describing a view of norms[93] that greatly influences my own, Witt (2011) writes, "feminist politics should be directed toward changing social reality rather than on changing the individual social agent and her choices" (47). Witt's primary insight here is that one way of changing the ways we think and interact with one another is through changing the norms that govern these behaviors. Though I do think the personal work required for sustaining oneself in long-term social change[94] is lost in Witt's articulation, what Witt zeros in on here is that personal commitments to changing the ways in which we interact with each other will not be enough

---

[92] To be clear, the norms that govern our epistemic practices lead to epistemic systems that are worse off epistemically and morally. 'Inadequacy' is meant to capture the missing epistemic resources and 'unjust' is meant to capture the ways in which this exclusion is immoral. In addition, these inadequacies lead to further moral problems because they lead to and support the Knowledge Problem, which is an epistemic barrier toward acting morally.

[93] I do not discuss Witt in my section on norms because her work is not situated within the norms debate, per say, though she has a lot to say about them. For more insight into how her ascriptivist view of norms can be used alongside the cognitive evolutionary account described by Kelly and Davis (2018), see Davidson and Kelly (2018).

[94] See Davidson and Gruver (2019) for more insight.

because these ways of interacting are grounded in collective practices and stabilized by internalized norms with which we are intrinsically motivated to comply and enforce on others. When we direct our energy toward changing social reality, our energy is focused on replacing unjust internalized norms with just norms. This is also true when we aim to adjust the norms that govern epistemic practices.

Kelly (forthcoming) develops a useful distinction between avowed norms, those norms that we actively endorse and wish to follow, and internalized norms, those norms that have been picked up by the norms system such that we have intrinsic motivation to comply to them. What Kelly, Witt, and I are all interested in, then, is how to internalize avowed norms such that they replace the unjust norms that we have already internalized, those norms we often follow when we don't want to be following them. But, this is difficult. We can't just force an avowed norm into our norm system. This is because our norm system picks up on norms through social learning, not through endorsement or route memorization. In this section, I'll utilize a third classification of norms, *candidate* norms, or those avowed norms we wish to "install" into a local network and are taking concrete steps to do so.

Because the norm system works through social learning, picking up on patterns of behavior and supplying intrinsic motivation to fall in line with these patterns, candidate norms often require that formal structures be put in place during the incubation process. These formal structures enforce the patterns of behavior in the absence of intrinsic motivation. This is needed because our dominant epistemic practices, governed by the norms we have already internalized, are likely to take over, especially when tensions arise or when one is moving quickly.

In this section, I'll detail one overarching candidate norm and three specific candidate norms, as well as the formal structures that allow them to incubate, that are being installed in several activist and organizing communities of which I'm a part.[95] These candidate norms serve to break down the norms that support harmful, dominant epistemic practices and replace them with

---

[95] An important thing to note about these organizations is that they see the anti-racist struggle as intimately tied to the struggles against other forms of oppression, including ableism, homophobia, transphobia, sexism, fatphobia, etc. Thus, when we talk about standing in solidarity with workers, we see this is as anti-racist since systems of oppression that oppress workers work in tandem with systems that oppress People of Color. As I discuss in the Introduction to the dissertation, my singular focus on racism particularly in the Chapters 2 and 3 is intentional, but does not include the robust intersectional analysis that would be required to understand systems of oppression as a whole or to understand individual experiences of oppression.

new norms that support just epistemic practices that will also address the Knowledge Problem. In short, these candidate norms are liberatory norms.

5.5.1 Process-Centered Candidate Norms

Imagine you and a friend have bought a model airplane and plan to put it together the next morning. After your friend has gone home for the evening, you start to think about the plane. You think about how difficult it will be to put the plane together, how slow the process will be, and how the plane probably won't come out looking exactly as it does on the box. You start to think that your friend really doesn't want to put together the plane after all and just wants the finished plane. After thinking about it for an hour or so, you decide to put the plane together yourself. You spend several hours on the plane and go to bed satisfied. When your friend comes over the next morning, they are devastated. You say to your friend, "Look how nice the plane is! I bet it can even fly." And your friend responds, "But I wanted to build it *together*."

One of the primary tenets in organizing groups of which I'm a part is that "the process is the product" and that the process usually involves doing things together. At times, it feels easier, more efficient, and quicker to do things on our own, in the way that feels most comfortable and natural to us, but, just like with the model airplane, doing things alone often misses the true point of working together to create sustainable change. The work itself, working across difference, paying attention to the patterns of social power, and building authentic relationships, is in itself anti-racist work.

Developing and practicing a process-centered model is an overarching candidate norm that, once installed, will influence the development and enforcement of new candidate norms. In this way, the process-centered norm is a meta-norm. Due to its nature, it is difficult to identify formal structures that will support the incubation of this candidate norm, but it is in many ways the most important candidate norm. This is because the processes we use to make our decisions will influence the decisions we make and how these decisions will be received by others.

One look at the structure of election systems will show you that this is the case: some ways of splitting populations into districts and counting their votes will result in multi-party systems and others result in two-party systems. That the process influences the outcome is the same within organizing groups. If an organization uses a "majority rules" voting system, then close to half of their members could disagree with a decision. If, on the other hand, an organization uses a

consensus model, then everyone must agree or have no strong disagreements with the decision before moving forward. Majority rules is certainly faster, but in focusing on the product (the decision) rather than focusing on the process of consensus building, it has the power to silence those who are oppressed in the current relations of power, especially when they are in the minority numbers-wise.

The commitment to being process-centered is also a radical recognition that our current practices, both epistemic and material, serve to uphold current relations of power and to maintain current epistemic practices. One of the commitments of our current processes is that the product matters more than the process (e.g. being a "good" white person justifies doing nothing about racism, increased capital justifies low wages and poor working conditions, "safety" justifies mass incarceration, border security justifies modern concentration camps). Thus, any candidate norm that brings to the fore the process will challenge this implicit commitment to outcomes. This is an important epistemic intervention because it is within the process or practice of our epistemic lives that the testimony of People of Color is excluded and that our inadequate epistemic systems survive.

The example of majority rules versus consensus above is easily codified. An organization can choose a model of decision-making and put it in their governing documents. This is a formal structure that supports the incubation of the process-centered candidate norm. However, because our dominant epistemic practices push us toward finishing, toward getting to the product (the decision, the next steps, the action items, etc.), the focus on the process rather than the product is hard shift to maintain in cases not governed by the formal decision-making process, the formal structure put in place to incubate the process-centered candidate norm. So, I'll put forward three further candidate norms that work to break down dominant epistemic practices in a way that improves our collective epistemic resources so that we can understand the experiences of more people. This allows us to address the Knowledge Problem, to more accurately identify instances of racism and use the intervention-sensitive model of responsibility to intervene on racist entities.

5.5.2   Centering those Most Affected

The second candidate norm, centering those most effected, is deeply rooted in a commitment to honoring epistemic advantage as it relates to social identities. This candidate norm has been developed from the in the ideas in bell hooks' ([1984] 2000) *Feminist Theory: From Margin to Center* in which she argues that Black women's lives are at the margins and hidden

from those working in mainstream feminist theory. As a result, feminist theory and practice, which was itself a disruption of dominant epistemic practices, will lack the epistemic resources for understanding the experiences of all women and will instead focus on white women and their struggles within a racially privileged identity. Thus, to develop adequate epistemic resources, those on the margins must be centered.

Within movement (activism and organizing) work, this candidate norm shows up in making sure that whenever we work on a particular issue, those who are most affected by that issue are front and center in the thinking and decision making around a particular campaign. For example, my organization recently stood in solidarity with over six-hundred janitors in Indianapolis, many of whom were Spanish-speaking Women of Color, in their fight for a fair contract for over a full year of negotiations. Since we did not have anyone in our organization that was a janitor in this particular union, we attended meetings to identify strategic ways to show our support (which often consisted in turning out people to events to show the companies that the janitors had power) rather than to decide how the campaign would look like or how they should move forward. Centering those most affected by an issue doesn't require that you do nothing, but rather requires listening to those who have knowledge about the issue (e.g. working conditions, needs of employees, etc.).

Centering those most affected by the issue is an ongoing practice, and it is often difficult to make sure that dominant epistemic norms do not creep in to the govern the situation. Incubating candidate norms is a bit like trying to rebuild a boat while you are still out at sea. For example, a few directors of my organization were together in an unofficial capacity when we began to discuss childcare at our meetings and whether our current offerings were serving mothers. After talking about it for several minutes, we realized that none of the people in the room were mothers. And, yet, we were discussing the issue as if we could have adequate insight into the effectiveness our childcare offerings. Rather than continue the conversation and make decisions on behalf of the mothers in our organization, we called on a few mothers to organize the other mothers for a meeting to discuss childcare options.

In some ways, centering those most affected by an issue is in line with well-known and endorsed epistemic ideals: those who have the most evidence should contribute the most to a conversation about an issue. Just like I would step back and allow a physicist to explain why the picture of the Black hole was so fuzzy, our collective epistemic practices should be in line with the ideals that experts should inform our decision making. However, because those who are experts

are experts because of a facet of oppression (e.g. People of Color are experts about the experience of racism because they experience it, folks without access to medical care are experts about the effects this has on them and their families, disabled individuals are experts about the ways in which spaces are inaccessible), our epistemic practices systematically exclude them as experts because including them would require a radical shift in epistemic practices (Hill Collins 1990).

Though this candidate norm can be supported by a formal structure when it is encoded into the policies and the expressed mission of movement organizations, this candidate norm requires near-constant vigilance because it is so often relevant within social change work. This means that replacing old norms around who knows and who gets to speak and installing candidate norms is integral to this process. One way of thinking about this is using Bicchieri's (2017) framework.[96] In order to shift this norm, we must change our empirical expectations. Rather believing that others will speak as though they were experts on every issue, especially when they occupy dominant positions, we should expect that others will speak only when they are experts and will amplify the voices of those who are experts with respect to a particular issue when they can. Further, and this is where an explicit mission is helpful, we can change our normative expectation. Rather than believing that those who are in the most dominant positions should offer their perspectives, we believe that they should not and that they should make room for others to have their voices heard. Once these changes in expectations have been made, the group-level epistemic practices begin to shift. As the norms shift, our epistemic practices change such that the voices of People of Color are amplified with respect to the experience of racism and the locations for intervention, addressing the Knowledge Problem.

5.5.3   Progressive Stack[97]

Even if we know it by another name, we're all somewhat accustomed to "taking stack" in question and answer sessions, workshops, and the classroom. This consists in paying attention to the (rough) order in which hands were raised, perhaps jotting it down, and calling on people in this

---

[96] In line with Bicchieri's (2017) framework, I've used "believe" here to indicate the changes that have occurred. She doesn't discuss it in her book, but I think the best interpretation is that these beliefs do not need to be conscious or reportable, but rather behavior-guiding (a distinct departure from how epistemologists use 'belief', but in line with the way many philosophers of mind use 'belief'). For a nice overview of the ways philosophers use 'belief', see Brownstein (2018, 72–81).

[97] I also write about progressive stack in a book chapter, "Epistemic Responsibility and Implicit Bias," co-authored with Nancy McHugh scheduled to appear in *Introduction to Implicit Bias*, edited by Alex Madva and Erin Beeghly.

order. Taking stack is by itself an intervention into epistemic practices governed by norms. In situations where we don't take stack—in everyday conversations, corporate meetings, or police investigations—norms that govern assignments of credibility and the identification of reliable knowers are guiding our behaviors. This means that those in dominant knowledge positions or those who are willing to speak loudly or interrupt (often overlapping groups) are those who are heard. Because taking stack also communicates that others should not be speaking once someone is called upon, stack gives the space for whomever wanted to be on the list to speak.

In addition to this form of taking stack, we might be familiar with some versions of revising the stack. For example, in my department, graduate students ask the first questions at colloquia. This version of revising the stack makes sure that growing scholars have the opportunity to ask questions before seasoned vets. Alternatively, at a conference, we might say that those who haven't yet asked a question will be bumped to the top of stack or that someone with a follow up may jump the stack by raising a finger to indicate that their question is a follow-up rather than a full-blown question. These revisions to the stack reflect commitments to specific kind of knowledge practices that disrupt dominant power relations in epistemic processes.

Progressive stack is also a revision to the stack, which serves as a formal structure to support the incubation of a candidate norm that disrupts dominant knowledge practices. For progressive stack, the person taking stack revises the stack order so that those who are most marginalized within current epistemic practices, those who are most often silenced, ignored, or otherwise epistemically oppressed, are moved to the top of stack where those who are most privileged within the current practices are moved to the bottom of the stack. In addition, this form of stack revision will likely be affected by the candidate norm of centering those most affected in that the stack will be revised such that those most affected by an issue will be able to speak toward the top of the stack.

Because progressive stack serves as a major intervention into dominant epistemic practices, practical issues may arise. For example, there will be times when engaging in progressive stack that not everyone will get to speak. And usually those who do not get to speak are those who are most used to getting their voices heard. This can result in backlash and strong emotions. If possible,

it is useful to have folks with similar identities[98] around to offer support after a meeting. In addition, is essential that the person who is taking stack is familiar with the members of the group because not all marginalization is visible. In cases where this is not possible, like large public meetings, folks may be asked to identify themselves as marginalized with respect to a particular issue. So, for example, one might say, "those who do not have adequate access to transportation will speak first, and those who do have access to transportation but would like to speak to this issue will speak second" before welcome people to a microphone.

Progressive stack is a formal structure that supports the incubation of candidate norms that govern epistemic practices. It works by disrupting our norm governed practices of deciding who gets to speak and for how long. Progressive stack is a policy intervention that serves to change norms over time within the local reference network that has adopted it. For example, the group I'm involved in doesn't always use progressive stack in a formal capacity, but the practice has shifted the norms that govern our epistemic practices even when we are not using the practice explicitly. In addition, it has continued to shift our collective hermeneutical resources because testimonies of those that may not have been heard given dominant epistemic practices are given space. The systematic inclusion of these voices means that epistemic resources are shared and new ones are developed for understanding the experiences of a wider variety of people and specifically of People of Color with respect to experiences of racism.

5.5.4   The Circle Way

One formal structure that helps to incubate candidate norms that organizations with which I am involved use is *The Circle Way*, developed by Baldwin and Linnea (2010). *Circle Way* is a structure for conversation that is an intervention into dominant epistemic practices because it interrupts several ways in which we communicate with each other and disrupts the ways communication can be used as a form of control. For example, within a Western context, we often use questions as a means of control. We ask questions that we already know the answers to or questions we want a specific answer to in order to control our interlocuter (this should be all too familiar to philosophers who participate in question and answer sessions). This kind of questioning is not allowed by the structure of discussion within the model of *The Circle Way*.

---

[98] Depending on the group, it may be white men that don't get a chance to speak. Rather than putting People of Color in the position of supporting the feelings of these individuals, it is best to identify other white men or woman that can provide this support.

The Circle Way is organized around three principles and three practices. The principles are rotating leadership, sharing responsibility and reliance on wholeness, and the three practice are attentive listening, intentional speaking, and attending to the wellbeing of the group (Baldwin and Linnea 2010, 26-28). When attending to the wellbeing of the group, one might ask oneself, "How do I offer my contribution in a way that will benefit what we are doing? (28). There are three forms of council offered by the practice, but I will focus on *talking piece council* as the type of council that moves away from dominant epistemic practices the most sharply. During this type of council, the group passes a talking piece, which can be anything, and during this time only the person with the talking piece can speak. This allows the speaker to contribute without fear of being interrupted. During this process, other members of the group attentively listen to the contribution, and the speaker focuses their contributions on what will best serve the group.

The Circle Way is not the only group- and process-centered structure for conversation; there are other ways. But, it is an effective formal structure that actively works against the dominant epistemic norms of conversations and serves to incubate candidate epistemic norms, so I've decided to use it as my example. Often times, conversations using this practice are met with resistance, and this is because so many of the structures put in place pull on our inner sense of how conversations should go, the structures pull on our internalized epistemic norms. After one practices *The Circle Way* for some time, this pull begins to fade. This is because the norm system has picked up on the practices of this new structure, this new way of having conversations. In my personal experience, a shift begins to happen such that you bring the norms of the practice into other conversations, which may result in social sanctions from those who have not experienced the norms that govern this process-centered epistemic practice.

5.5.5   Reference Networks, Trendsetters, and Installing Norms

I have just described four candidate norms and the formal structures that help to incubate these norms so that internalized norms don't take over. These liberatory epistemic norms have the power to change our ways of thinking and interacting such that the Knowledge Problem can be addressed, and we can build a liberated future. However, these norms are local norms. If we take seriously Bicchieri's (2017) notion of a reference network, or those people who are relevant to an individual when they are following norms, it stands to reason that a change in a particular reference network will have little to no effect on other reference networks. Norms that we want to cultivate

may get stuck in particular reference networks and will be unable to travel to those outside of a specific group. There's much work to be done to think this through completely, but looking to Bicchieri's notion of trendsetters (2017, 163-207) might push us in the right direction.

When a norm starts to spread through a particular reference network, *trendsetters* are the early adopters of a norm that pick up the norm and allow it to spread throughout the network. These trendsetters must have some specific characteristics in order for them to play this particular role. According to Bicchieri's behavioral analyses, trendsetters with respect to a particular norm are those with low personal allegiance to the norm (norm sensitivity) and low risk sensitivity in general and with respect to the particular norm (177). In addition, Bicchieri theorizes that trendsetters have in general a low sensitivity to the pressure to conform and a high level of perceived self-efficacy (167-172).

Even with the interventions to epistemic norms I've discussed above, trendsetters will be required. A new or replacement norm can't even get off the ground with an initial set of early adopters. However, in the case of these organizing spaces, people came together because they were seeking an intervention, seeking change to the dominant epistemic practices. Thus, the larger puzzle is not how these interventions into norms can get off the ground within an organizing context, but how these norms can begin to spread to the larger reference network in which these counter-cultures are situated.

In some ways, Bicchieri's methods for identifying trendsetters aren't compatible with the practices that are current underway in a United States context. Of course, she's writing in the context of non-governmental organizations, like UNICEF, making health interventions into communities that are very different from the communities the workers come from and different from each other. This differs from the context I'm thinking about here because I'm thinking about a small group (that constitutes a concentrated reference network) inside of a larger reference network. For example, she suggests surveying potential trendsetters to see how much their behavior changes in light of changing situations in order to measure risk sensitivity. In this case, one can identify low risk sensitivity when an individual's behaviors change when presented with new scenarios (i.e. they don't conform to the current norm when things change). This happens in organizing, though more informally. But, rather than helping to move the norm from the small group to the larger group, usually the identified trendsetter joins the smaller group.

Thus, the spread of norms from an organizing context to the larger society through trendsetters is likely to occur when members of the smaller reference network are interacting with those who are not a part of the smaller reference network at work, at school, or when socializing. In this way, they become infiltrators or undercover norm changers rather than trendsetters. Because of the new norms they have experienced within a particular reference network and the support they are getting from this network, they may feel more autonomous (a key feature of a good trendsetter) and be more likely to suggest changes or give feedback on epistemic practices that are upholding current material and epistemic conditions. Further, typically individuals who are covertly working against the dominant practices in their workplace, for example, are able to identify others who are doing the same and build coalitions within a workplace. This can lead to slow, but steady, epistemic norm shift over time.

5.6   Testimony in the Face of Death

In this chapter I've focused on the ways in which shifting the norms that govern our epistemic practices allows us to re-shape our world toward liberation and away from racism. One reason for this is the current dominant epistemic practices persistently exclude the testimony of People of Color, as well as the resources developed by People of Color for understanding the experience of racism. This means that within the scope of these practices, the resources available for identifying instances of racism are lacking, leading to the Knowledge Problem for moral responsibility. It is my hope that the candidate norms I've suggested here can continue to be developed in anti-racist communities so that the voices of racialized persons can be heard, and the knowledge can be integrated into our practices. This is a long-term vision. But learning to hear People of Color also has immediate effects.

A testimony that goes unheard is the last thing in the air before police kill People of Color. Here are just a few of the very many such testimonies: "I can't breathe." (Eric Garner), "You promised you wouldn't kill me." (Natasha McKenna), "Security" (Jemel Roberson), "I was reaching for…" (Philandro Castile), "I'm pregnant." (Name Unreleased). Epistemic practices matter in the moments between life and potential death in police interactions with People of Color. There are many psychological mechanisms at play in these moments, but one prominent mechanism are the norms that govern dominant epistemic practices, norms that lead people to

judge that Black and brown people are not trustworthy testifiers and that they have good reasons to mislead.

Movements like Black Lives Matter have demonstrated the flaws in the dominant epistemic practices by demanding to be heard, by showing their power through marches and other demonstrations. In articulating a set of demands that show the multi-dimensional ways in which there is an active assault on Black bodies in this country and throughout the world, they intervene on our collective epistemic resources by producing a framework for understanding experiences of racism. Epistemic norm interventions at the interpersonal, cultural, and structural levels will contribute to a robust anti-racist project and have the potential to combat racism, to prevent suffering and pre-mature death.

CHAPTER 6.   CONCLUSION

I started this dissertation by highlighting philosophers who suggest that we limit our use of the terms 'racism' and 'racist'. This suggestion is motivated by a worry that the concept RACISM has become incoherent, lumpy, and inflated such that our uses no longer refer to anything meaningful. Due to these conceptual worries, they also fear that our ability to morally evaluate one another, to hold each other accountable, with these terms is diminished. People outside academy have expressed similar worries. Joking uses of "that's racist!" point to the general feeling that terms 'racism' and 'racist' are overused. In response, I've laid out a project that increases the conceptual clarity of RACISM, specifically with respect to understanding locations of agreement and disagreement within the individualist versus structuralist debate and when and why we should apply the predicate 'racist' to particular entities. From this clarity, I've then laid out a new framework for understanding moral responsibility for racism and the liberatory epistemic norms that must be installed for this responsibility to be enacted.

I've done so by developing a new frame, levels of pluralism, for understanding the commitments different accounts of racism take on, explicitly or implicitly. Within this framework, I've identified a set of four nested questions that are in a dependence relationship with one another. The outcome is that I've rejected entity type monism, or the view that only certain kinds of entities are the fundamental bearers of the predicate 'racist', which is endorsed by both individualists and structuralists. This implicit endorsement arises due to a failure to recognize or appropriately appreciate the dependence relationship between the nested questions. Reject of entity type monism clears the ground for discussions of racism that go beyond the individualist versus structuralist divide and increases clarity around the concept RACISM.

From the rejection of entity type monism, I've developed an account of the application conditions for the predicate 'racist' that shows the predicate can be appropriately applied to many different entities. The predicate is appropriately applied to (A) an entity (B) when the entity plays some role in an individual suffering harms (C) that prematurely kill and prevent from being born and (D) when those harms are: (D.1) unjustified, (D.2) suffered due to race, (D.3) and conditionally probable to be suffered given race in a given context. This articulation is pluralist with respect to entity type, causation, and intervention strategies. Unlike conversations within the individualist

versus structuralist debate, this leaves open the possibility for many causal mechanisms and many locations for intervention with respect to racism.

The first concern articulated by those who seek to limit the use of 'racism' and 'racist' is that the concepts are incoherent or inflated. The second concern is that this lack of clarity means that our attempts to morally evaluate one another using these terms fail. After increasing the clarity of RACISM in the Chapters 2 and 3, I developed new resources for understanding moral evaluation with respect to racism, oblique blame and intervention-sensitive moral responsibility. These two concepts allow individuals to be held responsible in their relationships to non-agential racist entities.

When a person (A) obliquely blames another person (T), A judges that an entity (E) is racist, A thinks that T is related to E such that their intervention may be effective, and A calls on T to intervene on E by directing a response or appraisal toward T. Effectively, oblique blame is a call to action. Out of oblique blame, I developed a concept of intervention-sensitive moral responsibility: T should intervene on E when E is racist (or has some other negative moral property) and T is related to E in a way such that T's intervention may be effective. These conceptions neatly avoid problems of attributability and the related strategies for avoiding responsibility for racism.

From oblique responsibility and intervention-sensitive responsibility, a problem arises: how are we to know when we are responsible in this intervention-sensitive sense when the dominant epistemic practices serve to obscure matters related to race (Mills 1997)? Further, how can oblique blame be effective given the lower credibility often assigned to People of Color, those who are in the best position to judge when an entity is racist (a condition for oblique blame) given their experiences with racism? I've called this the Knowledge Problem for Intervention-Sensitive Moral Responsibility.

I've united the literature on epistemic systems and oppression with the empirically-informed literature on norms to show that norms are an effective location for addressing the Knowledge Problem. I explored four candidate social norms, supported by formal structures during the incubation process, that serve to disrupt dominant epistemic norms in ways that transform our epistemic practices. These norms are focusing on the process rather than the product, centering those most effected by an issue, progressive stacking, and *The Circle Way* (a set of practices that support non-hierarchical, shared leadership practices within conversation). The

norms are liberatory in that they work together to change the epistemic landscape in ways that make holding each other accountable for racism more successful.

Through this project, I've shown that gaining conceptual clarity around the concept RACISM, specifically the predicate 'racist', is possible and that doing so helps us identify pathways for sustainable change. I've done so because abandoning these terms and concepts— terms and concepts that have and continue to transform the racial landscape and build a more liberated future—is antithetical to the anti-racist project. Given my aims, here are some areas for future research.

6.1 Moral Responsibility

As I've articulated it in Chapter 4, oblique blame runs afoul of a major tenet of the moral responsibility literature: blame is never appropriate without attributability. However, if you look at our practices of moral responsibility and blame for racism, people feel blamed without attributability because 'racist' is a thick moral description. I'd like to look deeper into the theoretical origin of the tight coupling of blame and attributability to analyze its foundations and look more deeply for resources that analyze feelings of guilt, shame, or otherwise being blamed within conversations of race and racism. Also in Chapter 4, I mentioned that oblique blame is directed at individual agents, but that I would like to explore the ways in which oblique blame can be directed at collective agents. Specifically, I'd like to bring together the philosophical literature on moral responsibility with the work Guala (2016) has done to articulate different facets of institutions.

6.2 Accountability and "Everyday" Communication

In continuing to look to ordinary practices of blame, I'd like to analyze the subversive accountability strategies taken up by the left, particularly on the internet. For example, "all men are trash" is often used to point out sexist entities, but it was recently banned by Facebook as hate speech because men felt unduly targeted or blamed. This phrase functions similarly to "all white people are racist," which as I explored in Chapter 4, garners negative reactions even when no one said it. And, recently I've been enjoying frantic reactions to "White women killed yoga." In future research, I'd like to explore what is expressed by these utterances and what is communicated. I

think these should be interpreted with the literature on generics and communicative failures, but any final conclusions will need to wait until the research is done.

Also related to communication and accountability, many people seem to think that what they *have* said is what they *meant* to say. Specifically, they think if they didn't *mean* to say something racist, they didn't. I'm interested in analyzing why reactions like, "I didn't mean it that way," fail in light of an externalist view of meaning.

6.3   Installing Norms and Rebuilding Trust

In Chapter 5, I began to explore the ways in which local norms, or those norms that have been installed within a particular reference network, might be spread more globally. I utilized Bicchieri's (2017) notion of trendsetters to show how individual early adopters may be integral to this process. I'd like to continue to think more deeply about how this process may work, including looking at the empirical literature in behavioral economics and the concept of norm entrepreneurs developed in political science.

In another branch of my research, I explore the ways in which oppressed individuals and groups are justified in global social distrust, or distrust of many of the members in their community and the institutions within it. Within the public health context, I argue that health providers have a responsibility to re-build trust with the communities that are currently failing to serve. Rebuilding trust requires both changing the institutions themselves so they are trustworthy and developing relationships to repair past damage and address the affective side of distrust. Thus far, I have not analyzed any specific strategies for rebuilding trust, and I'd like to think more about whether or not the strategies I discuss in Chapter 5 can be applied to this context.

All told, I reckon these projects could take up the rest of my career as a philosopher, so I won't list any more of my interests. The overarching theme of these future projects, as well as the project of the dissertation, is the view that philosophy can be used as a tool to think carefully about the social world, the ways in which people die, suffer, and are unjustly constrained and also the possibilities for freedom and liberation, and the view that thinking about the social world will provide us with new insights about the best pathways for creating sustained positive change for a more liberated future.

REFERENCES

Acciai, Francesco, Aggie J Noah, and Glenn Firebaugh. 2015. "Pinpointing the Sources of the Asian Mortality Advantage in the USA." *Journal of Epidemiology and Community Health* 69 (10): 1006–11. https://doi.org/10.1136/jech-2015-205623.

Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2018. *Deep Roots: How Slavery Still Shapes Southern Politics*. Princeton: Princeton University Press.

Alexander, Michelle, and Cornel West. 2012. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. Revised edition. New York: New Press.

Allport, Gordon. 1954. *The Nature of Prejudice*. Cambridge, MA: Addison-Wesley.

Anderson, Elizabeth. 2010. *The Imperative of Integration*. Princeton, NJ: Princeton University Press.

Appiah, Kwame A. 1990. "Racisms." In *Anatomy of Racism*, edited by David Theo Goldberg. Minneapolis: University of Minnesota Press.

Appiah, Kwame A. 1992. *In My Father's House: Africa in the Philosophy of Culture*. New York: Oxford University Press.

Ashton, Carol M., Paul Haidet, Debora A. Paterniti, Tracie C. Collins, Howard S. Gordon, Kimberly O'Malley, Laura A. Petersen, et al. 2003. "Racial and Ethnic Disparities in the Use of Health Services: Bias, Preferences, or Poor Communication?" *Journal of General Internal Medicine* 18 (2): 146–52. https://doi.org/10.1046/j.1525-1497.2003.20532.x.

Baldwin, Christina, and Ann Linnea. 2010. *The Circle Way: A Leader in Every Chair*. 1st ed. A BK Business Book. San Francisco, CA: Berrett-Koehler Publishers.

Barker, Simon, Charlie Crerar, and Trystan S. Goetze. 2018. "Harms and Wrongs in Epistemic Practice." *Royal Institute of Philosophy Supplement* 84 (November): 1–21. https://doi.org/10.1017/S1358246118000528.

Bechtel, William, and Adele Abrahamsen. 2005. "Explanation: A Mechanist Alternative." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 36 (2): 421–41. https://doi.org/10.1016/j.shpsc.2005.03.010.

Bertrand, Marianne, Dolly Chugh, and Sendhil Mullainathan. 2005. "Implicit Discrimination." *The American Economic Review* 95 (2): 94–98.

Bicchieri, Cristina. 2006. *The Grammar of Society: The Nature and Dynamics of Social Norms*. New York: Cambridge University Press.

Bicchieri, Cristina. 2017. *Norms in the Wild: How to Diagnose, Measure, and Change Social Norms*. New York, NY: Oxford University Press.

Blum, Lawrence. 2002. "Racism: What It Is and What It Isn't." *Studies in Philosophy and Education* 21: 203–18. https://doi.org/10.1007/978-1-4020-8626-7_6.

Blum, Lawrence. 2007. *"I'm Not a Racist, but ...": The Moral Quandary of Race*. Nachdr. Ithaca: Cornell Univ. Press.

Blum, Lawrence A. 2010. "Racialized Groups: The Sociohistorical Consensus." *The Monist* 93 (2): 298–320.

Blum, Lawrence. 2014. "What Do Accounts of 'Racism' Do?" In *Racism in Mind*, edited by Mark Levine and Tamas Pataki. Ithaca, New York: Cornell University Press.

Bonilla-Silva, Eduardo. 2003. *Racism Without Racists*. Lanham, Maryland: Rowman and Littlefield.

Bougie, Olga, Jenna Healey, and Sukhbir S. Singh. 2019. "Behind the Times: Revisiting Endometriosis and Race." *American Journal of Obstetrics and Gynecology*, February, S0002937819303229. https://doi.org/10.1016/j.ajog.2019.01.238.

Brown, Jessica. 2012. "Assertion and Practical Reasoning: Common or Divergent Epistemic Standards?" *Philosophy and Phenomenological Research* 84 (1): 123–57. https://doi.org/10.1111/j.1933-1592.2010.00432.x.

Brownstein, Michael. 2016. "Attributionism and Moral Responsibility for Implicit Bias." *Review of Philosophy and Psychology* 7 (4): 765–86. https://doi.org/10.1007/s13164-015-0287-7.

Brownstein, Michael. 2018. *The Implicit Mind: Cognitive Architecture, the Self, and Ethics*. New York, NY: Oxford University Press.

Buck v. Bell. 1927. U.S. Supreme Court.

Burgess, Alexis, and David Plunkett. 2013a. "Conceptual Ethics I." *Philosophy Compass* 8 (12): 1091–1101. https://doi.org/10.1111/phc3.12086.

Burgess, Alexis, and David Plunkett. 2013b. "Conceptual Ethics II." *Philosophy Compass* 8 (12): 1102–10. https://doi.org/10.1111/phc3.12085.

Cameron, C. Daryl, B. Keith Payne, and Joshua Knobe. 2010. "Do Theories of Implicit Race Bias Change Moral Judgments?" *Social Justice Research* 23 (4): 272–89. https://doi.org/10.1007/s11211-010-0118-z.

Carlsson, Andres Brekke. 2017. "Blameworthiness as Deserved Guilt." *Ethics* 21: 89–115.

Carmichael, Stokely, and Charles V. Hamilton. 1967. *Black Power*. New York, NY: Random House.

Caro, Robert A. 1975. *The Power Broker: Robert Moses and the Fall of New York*. New York: Vintage Books.

Carter, Evelyn R., and Mary C. Murphy. 2015. "Group-Based Differences in Perceptions of Racism: What Counts, to Whom, and Why?: Perceptions of Racism." *Social and Personality Psychology Compass* 9 (6): 269–80. https://doi.org/10.1111/spc3.12181.

Cashmore, Ernest. 1987. *The Logic of Racism*. London, UK: Allen & Unwin.

Center for Disease Control and Prevention/CDC WONDER. 2016. *Mapping of American 2018 (new), Health*. Distributed by Measure of America. http://measureofamerica.org/maps/

Chappell, Bill. 2013. "California's Prison Sterilzations Reportedly Echo Eugenics Era." *NPR*, July 9, 2013. https://www.npr.org/sections/thetwo-way/2013/07/09/200444613/californias-prison-sterilizations-reportedly-echoes-eugenics-era.

Chatman, Donald L. 1976. "Endometriosis in the Black Woman." *American Journal of Obstetrics & Gynecology* 125 (7): 987–89. https://doi.org/10.1016/0002-9378(76)90502-0.

Coates, D. Justin, and Neal A. Tognazzini. 2013. "The Contours of Blame." In *Blame: Its Nature and Norms*, edited by D. Justin Coates and Neal A. Tognazzini, 3–26. Oxford: Oxford University Press.

Coates, Ta-Nehisi. 2008. "Playing the Racist Card." *Slate News*, March 14, 2008. https://slate.com/news-and-politics/2008/03/ferraro-s-comments-about-obama-were-racist-why-can-t-we-say-that.html

"Criminal Justice Fact Sheet." n.d. NAACP. Accessed May 9, 2019. https://www.naacp.org/criminal-justice-fact-sheet/.

davenport, doris. [1981] 2015. "The Pathology of Racism: A Conversation with Third World Wimmin." In *This Bridge Called My Back* (4th ed., pp. 81–86). Albany: SUNY Press.

Davidson, Lacey J. 2017. "Category Matters: The Interlocking Epistemic and Moral Costs of Implicit Bias." *Teoria* 37 (2): 37-51. https://doi.org/10.4454/teoria.v37i2.16

Davidson, Lacey J., and Daniel Kelly. 2018. "Minding the Gap: Bias, Soft Structures, and the Double Life of Social Norms." *Journal of Applied Philosophy*, December. https://doi.org/10.1111/japp.12351.

Davidson, Lacey J., and Melissa D. Gruver. 2019. "Epistemic Labor and the Power in (Fat) Identity: Three Fat Archetypes as Experienced in the Local Bar, the Streets, and Other Public Places." *Fat Studies*. https://doi.org/10.1080/21604851.2019.1629806.

Davidson, Lacey J. manuscript. "When Testimony Isn't Enough: Implicit Bias Research as Epistemic Exclusion." In *Overcoming Epistemic Injustice*.

Davis, Taylor and Daniel Kelly. 2018. Norms, Not Moral Norms: The Boundaries of Morality Do Not Matter. *Behavioral and Brain Sciences* 41. https://doi-org/10.1017/S0140525X18000067

Delacampagne, Christian. 1983. *L'invention Du Racisme: Antiquité et Moyen Age*. Paris: Fayard.

Desmond, Matthew, and Mustafa Emirbayer. 2009. "What is Racial Domination?" *Du Bois Review* 6 (2): 335–355. doi:10.10170S1742058X09990166

DiAngelo, Robin. 2018. *White Fragility: Why It's So Hard for White People to Talk About Racism*. Boston, Massachusetts: Beacon Press.

Dotson, Kristie. 2011. "Tracking Epistemic Violence, Tracking Practices of Silencing." *Hypatia* 26 (2): 236–57. https://doi.org/10.1111/j.1527-2001.2011.01177.x.

Dotson, Kristie. 2012. "A Cautionary Tale: On Limiting Epistemic Oppression." *Frontiers: A Journal of Women Studies* 33 (1): 24. https://doi.org/10.5250/fronjwomestud.33.1.0024.

Dotson, Kristie. 2014. "Conceptualizing Epistemic Oppression." *Social Epistemology* 28 (2): 115–38. https://doi.org/10.1080/02691728.2013.782585.

Dunbar-Ortiz, Roxanne. 2014. *An Indigenous People's History of the United States*. Boston: Beacon Press.

Eshleman, Andrew. 2014. "Moral Responsibility." In *Stanford Encylopedia of Philosophy*. https://plato.stanford.edu/entries/moral-responsibility/.

Fabien, Vanessa. 2018. "My Body, My Pain: Listen to Me and All Black Women." The Root. Accessed October 4, 2018. https://www.theroot.com/my-body-my-pain-listen-to-me-and-all-Black-women-1794332651.

Fodor, Jerry A. 1990. *A Theory of Content and Other Essays*. Cambridge, Mass.: MIT Press.

Fodor, Jerry. 2004. "Having Concepts: A Brief Refutation of the Twentieth Century." *Mind and Language* 19 (1): 29–47. https://doi.org/10.1111/j.1468-0017.2004.00245.x.

Frankenberg, Ruth. 1993. *White women, race matters: the social construction of whiteness*. Minneapolis: University of Minnesota Press.

Fricker, Miranda. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford ; New York: Oxford University Press.

Frost, Mark B., Gerald J. Williams, Samuel Merovitz, and David M. Cedar. 2005. "Protecting the Right to Clean Living: In Poor and Minority Neighborhoods, Many Homes Are Tainted with Pollution. The Result: Illness and Declining Property Values. Here's How One Community Is Fighting Back." *Trial* 41 (2): 30–35.

Frye, Marilyn. 1983. *The Politics of Reality: Essays in Feminist Theory*. Berkeley, CA: Crossing Press.

Garcia, Jorge L.A. 1996. "The Heart of Racism." *Journal of Social Philosophy* 27 (1): 5–46. https://doi.org/10.1111/j.1467-9833.1996.tb00225.x.

Garcia, Jorge L.A. 2011. "Racism, Psychology, and Morality: Dialogue with Faucher and Machery." *Philosophy of the Social Sciences* 41 (2): 250–68. https://doi.org/10.1177/0048393109355268.

Gee, Gilbert C., Anna Hing, Selina Mohammed, Derrick C. Tabor, and David R. Williams. 2019. "Racism and the Life Course: Taking Time Seriously." *American Journal of Public Health* 109 (S1): S43–47. https://doi.org/10.2105/AJPH.2018.304766.

Gerken, Mikkel. 2014. "Same, Same but Different: The Epistemic Norms of Assertion, Action and Practical Reasoning." *Philosophical Studies* 168 (3): 725–44. https://doi.org/10.1007/s11098-013-0156-0.

Glasgow, Joshua. 2009. "Racism as Disrespect." *Ethics* 120 (1): 64-93.

Gordon, Lewis R. 1999. *Bad Faith and AntiBlack Racism*. New York, N.Y: Humanity Books.

Guala, Francesco. 2016. *Understanding Institutions*. Princeton: Princeton University Press.

Gupta, Anil. 2015. "Definitions." Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/definitions/

Hacking, Ian. 2001. *The Social Construction of What?* 8. printing. Cambridge, Mass: Harvard Univ. Press.

Harding, Sandra G. 1991. *Whose Science? Whose Knowledge? Thinking from Women's Lives*. Ithaca, N.Y: Cornell University Press.

Harris, Leonard. 1998. "The Concept of Racism: An Essentially Contested Concept? *The Centennial Review* 42 (2): 217-231

Harris, Leonard. 1999. "What, Then, Is Racism?" In *Racism*, edited by Leonard Harris, 437–50. Key Concepts in Critical Theory. Amherst, N.Y: Humanity Books.

Harris, Leonard 2018. "Necro-Being: An Actuarial Account of Racism." *Res Philosophica* 95 (2): 273–302. https://doi.org/10.11612/resphil.1617.

Harris, Ricci, Donna Cormack, Martin Tobias, Li-Chia Yeh, Natalie Talamaivao, Joanna Minster, and Roimata Timutimu. 2012. "The Pervasive Effects of Racism: Experiences of Racial Discrimination in New Zealand over Time and Associations with Multiple Health Domains." *Social Science & Medicine* 74 (3): 408–15. https://doi.org/10.1016/j.socscimed.2011.11.004.

Haslanger, Sally. 2012a. "Gender and Race: (What) Are They? (What) Do We Want Them To Be?" In *Resisting Reality: Social Construction and Social Critique*, 221–47. New York: Oxford University Press.

Haslanger, Sally. 2012b. "Language, Politics, and 'The Folk:' Looking for 'The Meaning' of 'Race.'" In *Resisting Reality: Social Construction and Social Critique*, 429–45. New York: Oxford University Press.

Haslanger, Sally. 2012c. *Resisting Reality: Social Construction and Social Critique*. New York: Oxford University Press.

Haslanger, Sally. 2015. "Distinguished Lecture: Social Structure, Narrative and Explanation." *Canadian Journal of Philosophy* 45 (1): 1–15. https://doi.org/10.1080/00455091.2015.1019176.

Haslanger, Sally. 2018. "What Is a Social Practice?" *Royal Institute of Philosophy Supplement* 82 (July): 231–47. https://doi.org/10.1017/S1358246118000085.

Henrich, Joseph. 2016. *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species and Making Us Smarter*. Princeton Oxford: Princeton University Press.

Heron, Melonie. 2017. "Deaths: Leading Causes for 2015." 66.5. Hyattsville, Maryland: National Vital Statistics Reports.

Hill Collins, Patricia. 1990. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. Perspectives on Gender, v. 2. Boston: Unwin Hyman.

Hoffman, Kelly M., Sophie Trawalter, Jordan R. Axt, and M. Norman Oliver. 2016. "Racial Bias in Pain Assessment and Treatment Recommendations, and False Beliefs about Biological Differences between Blacks and Whites." *Proceedings of the National Academy of Sciences* 113 (16): 4296–4301. https://doi.org/10.1073/pnas.1516047113.

Holroyd, Jules. 2012. "Responsibility for Implicit Bias." *Journal of Social Philosophy* 43 (3): 274–306. https://doi.org/10.1111/j.1467-9833.2012.01565.x.

hooks, bell. 2000. *Feminist Theory: From Margin to Center*. 2nd ed. South End Press Classics, v. 5. Cambridge, MA: South End Press.

Hozumi, Tada. 2017. "Setting Boundaries as a White Ally: Why Its Important, Why Its Challenging, and How to Do It Ethically." *Selfish Activist* (blog). July 5, 2017. https://selfishactivist.com/how-to-set-boundaries-as-a-white-ally/.

Hyland, Shelley, Lynn Langton, and Elizabeth Davis. 2015. "Police Use of Nonfatal Force, 2002-11." Special Report NCJ 249216. U.S. Department of Justice.

Jagger, Alison. 1988. *Feminist Politics and Human Nature*. Lanham: Rowman and Littlefield.

Kahneman, Daniel. 2013. *Thinking, Fast and Slow*. 1st pbk. ed. New York: Farrar, Straus and Giroux.

Kaiser Family Foundation. *Life Expectancy at Birth (in years), by Race/Ethnicity.* Distributed by the Kaiser Family Foundation. https://www.kff.org/other/state-indicator/life-expectancy-by-re/

Kauppinen, Antti. 2018. "Epistemic Norms and Epistemic Accountability." *Philosophers' Imprint* 18 (8): 1–16.

Kawall, Jason. 2013. "Friendship and Epistemic Norms." *Philosophical Studies* 165 (2): 349–70. https://doi.org/10.1007/s11098-012-9953-0.

Keller, Simon. 2004. "Friendship and Belief." *Philosophical Papers* 33 (3): 329–51. https://doi.org/10.1080/05568640409485146.

Kelly, Daniel, and Taylor Davis. 2018. "Social Norms and Human Normative Psychology," 19.

Kelly, Daniel. Forthcoming. "How to Adopt a Norm: Toward a Pluralist Account of the Psychology of Norms." In *The Oxford Handbook of Moral Psychology*, edited by Manuel Vargas and John Doris.

Khader, Serene J. 2011. *Adaptive Preferences and Women's Empowerment*. Studies in Feminist Philosophy. New York: Oxford University Press.

Knobe, Joshua, and John M. Doris. 2012. "Responsibility." In *The Moral Psychology Handbook*, edited by John M. Doris, 1st ed., 321–54. Oxford: Oxford Univ. Press.

Kovel, Joel. 1970. *White Racism: A Psychohistory*. New York, NY: Pantheon.

Kukla, André. 2000. *Social Constructivism and the Philosophy of Science*. Philosophical Issues in Science. London: Routledge.

Kyama, M. C., T. M. D'Hooghe, S. Debrock, J. Machoki, D. C. Chai, and J. M. Mwenda. 2004. "The Prevalence of Endometriosis among African-American and African-Indigenous Women." *Gynecologic and Obstetric Investigation* 57 (1): 40–42.

Levy, Neil. 2017. "Implicit Bias and Moral Responsibility: Probing the Data." *Philosophy and Phenomenological Research* 94 (1): 3–26. https://doi.org/10.1111/phpr.12352.

Lewis, Tené T., Courtney D. Cogburn, and David R. Williams. 2015. "Self-Reported Experiences of Discrimination and Health: Scientific Advances, Ongoing Controversies, and Emerging Issues." *Annual Review of Clinical Psychology* 11 (1): 407–40. https://doi.org/10.1146/annurev-clinpsy-032814-112728.

Linder, Chris. 2015. "Navigating Guilt, Shame, and Fear of Appearing Racist: A Conceptual Model of Antiracist White Feminist Identity Development." *Journal of College Student Development* 56 (6): 535–50. https://doi.org/10.1353/csd.2015.0057.

Link, Bruce G., and Jo C. Phelan. 1995. "Social Conditions as Fundamental Casues of Disease." *Journal of Health and Social Behavior* 35: 80–84.

Lugones, María. 1987. "Playfulness, 'World'-Traveling, and Loving Perception." *Hypatia* 2 (2): 3-19. https://doi.org/10.1111/j.1527-2001.1987.tb01062.x

Lycan, William G. 1995. *Consciousness*. 1. MIT Pr. paperback ed. A Bradford Book. Cambridge, Mass.: MIT Pr.

Machamer, Peter, Lindley Darden, and Carl F. Craver. 2000. "Thinking about Mechanisms." *Philosophy of Science* 67 (1): 1–25. https://doi.org/10.1086/392759.

Madva, Alex. 2016. "A Plea for Anti-Anti-Individualism: How Oversimple Psychology Misleads Social Policy." *Ergo, an Open Access Journal of Philosophy* 3 (20180912). https://doi.org/10.3998/ergo.12405314.0003.027.

Mallon, Ron. 2006. "'Race'': Normative, Not Metaphysical or Semantic.'" *Ethics* 116 (3): 525–51. https://doi.org/10.1086/500495.

Mallon, Ron. 2016. *The Construction of Human Kinds*. First edition. Oxford New York, NY: Oxford University Press.

Mallon, Ron. 2017. "Social Construction and Achieving Reference: Social Construction and Achieving Reference." *Noûs* 51 (1): 113–31. https://doi.org/10.1111/nous.12107.

Mason, Elinor. 2019. *Ways to Be Blameworthy: Rightness, Wrongness, and Responsibility*. New product edition. Newyork, NY: Oxford University Press.

Mbembe, Achille. 2003. "Necropolitics." *Public Culture* 15 (1): 11–40.

McHugh, Nancy, and Lacey Davidson. forthcoming. "Epistemic Responsibility and Implicit Bias." In *An Introduction to Implicit Bias: Knowledge, Justice, and the Social Mind*. New York, NY: Routledge.

McIntosh, Peggy. 1988. "White Privilege and Male Privilege: A Personal Account of Coming to See Correspondences through Work in Women's Studies." In *Race, Class, and Gender: An Anthology*, edited by M.L. Anderson and Patricia Hill Collins, 76–87. Belmont, CA: Wadsworth.

McKinnon, Rachel R. 2012. "What I Learned in the Lunch Room about Assertion and Practical Reasoning:" *Logos & Episteme* 3 (4): 565–69. https://doi.org/10.5840/logos-episteme2012343.

Medina, José. 2012. "Hermeneutical Injustice and Polyphonic Contextualism: Social Silences and Shared Hermeneutical Responsibilities." *Social Epistemology* 26 (2): 201–20. https://doi.org/10.1080/02691728.2011.652214.

Miles, Robert. [1989] 1993. *Racism*. London, UK: Routledge.

Miller, Ted R, Bruce A Lawrence, Nancy N Carlson, Delia Hendrie, Sean Randall, Ian R H Rockett, and Rebecca S Spicer. 2017. "Perils of Police Action: A Cautionary Tale from US Data Sets." *Injury Prevention* 23 (1): 27–32. https://doi.org/10.1136/injuryprev-2016-042023.

Mills, Charles W. 1997. *The Racial Contract*. Ithaca, NY: Cornell University Press.

Mills, Charles W. 1998. "Alternative Epistemologies." In *Epistemology: The Big Questions*, edited by Linda Martin Alcoff, 4th repr, 392–410. Malden: Blackwell.

Mills, Charles W. 2005. "'Ideal Theory' as Ideology." *Hypatia* 20 (3): 165–84.

Montminy, Martin. 2013. "Why Assertion and Practical Reasoning Must Be Governed By the Same Epistemic Norm: Assertion and Practical Reasoning." *Pacific Philosophical Quarterly* 94 (1): 57–68. https://doi.org/10.1111/j.1468-0114.2012.01444.x.

"Murder and Extremism in the United States in 2018." 2018. Anti-Defamation League. https://www.adl.org/murder-and-extremism-2018#the-perpetrators.

Musemwa, Nomsa, and Crystal A. Gadegbeku. 2017. "Hypertension in African Americans." *Current Cardiology Reports* 19 (12): 129. https://doi.org/10.1007/s11886-017-0933-z.

Nussbaum, Martha C. 1995. "Human Capabilities, Female Human Beings." In *Women, Culture, and Development*, edited by Jonathan Glover and Martha C. Nussbaum, 61–104. Oxford University Press. https://doi.org/10.1093/0198289642.001.0001.

O'Neill, Onora. 1975. "Lifeboat Earth." *Philosophy and Public Affairs* 4 (3): 273-292.

Oluo, Ijeoma. 2018. *So You Want to Talk about Race*. New York: Seal Press.

Omi, Michael, and Howard A. Winant. 2015. *Racial Formation in the United States*. Third edition. New York London: Routledge Taylor & Francis Group.

Onukwugha, Ebere, and C. Daniel Mullins. 2007. "Racial Differences in Hospital Discharge Disposition among Stroke Patients in Maryland." *Medical Decision Making* 27 (3): 233–42. https://doi.org/10.1177/0272989X07302130.

Outlaw, Lucius T. 1996. *On Race and Philosophy*. New York: Routledge.

Paradies, Yin. 2006. "A Systematic Review of Empirical Research on Self-Reported Racism and Health." *International Journal of Epidemiology* 35 (4): 888–901. https://doi.org/10.1093/ije/dyl056.

Phelan, Jo C., and Bruce G. Link. 2015. "Is Racism a Fundamental Cause of Inequalities in Health?" *Annual Review of Sociology* 41 (1): 311–30. https://doi.org/10.1146/annurev-soc-073014-112305.

Pletcher, Mark J., Stefan G. Kertesz, Michael A. Kohn, and Ralph Gonzales. 2008. "Trends in Opioid Prescribing by Race/Ethnicity for Patients Seeking Care in US Emergency Departments." *JAMA* 299 (1). https://doi.org/10.1001/jama.2007.64.

Plunkett, David, and Tim Sundell. 2013. "Disagreement and the Semantics of Normative and Evaluative Terms." *Philosopher's Imprint* 13 (23): 1–37.

Pohlhaus, Gaile. 2012. "Relational Knowing and Epistemic Injustice: Toward a Theory of Willful Hermeneutical Ignorance." *Hypatia* 27 (4): 715–35. https://doi.org/10.1111/j.1527-2001.2011.01222.x.

Portmore, Douglas. manuscript. "A Comprehensive Account of Blame: Self-Blame, Non-Moral Blame, and Blame for the Non-Agential."

Priest, Naomi, Yin Paradies, Brigid Trenerry, Mandy Truong, Saffron Karlsen, and Yvonne Kelly. 2013. "A Systematic Review of Studies Examining the Relationship between Reported Racism and Health and Wellbeing for Children and Young People." *Social Science & Medicine* 95 (October): 115–27. https://doi.org/10.1016/j.socscimed.2012.11.031.

Rasooly, Irit R., Peter M. Mullins, Maryann Mazer-Amirshahi, John van den Anker, and Jesse M. Pines. 2014. "The Impact of Race on Analgesia Use among Pediatric Emergency Department Patients." *The Journal of Pediatrics* 165 (3): 618–21. https://doi.org/10.1016/j.jpeds.2014.04.059.

Relf v. Weinberger. 1974. U.S.D.C. District of Columbia.

Richerson, Peter J., and Robert Boyd. 2008. *Not by Genes Alone: How Culture Transformed Human Evolution*. Paperback ed., [Nachdr.]. Chicago, Ill.: Univ. of Chicago Press.

Rolin, Kristina. 2009. "Standpoint Theory as a Methodology for the Study of Power Relations." *Hypatia* 24 (4): 218–26. https://doi.org/10.1111/j.1527-2001.2009.01070.x.

Ross, Don. 2002. "Why People Are Atypical Agents." *Philosophical Papers* 31 (1): 87–116. https://doi.org/10.1080/05568640209485097.

Ross, Loretta, and Rickie Solinger. 2017. *Reproductive Justice: An Introduction*. Reproductive Justice : A New Vision for the Twenty-First Century 1. Oakland, California: University of California Press.

Scanlon, Thomas. 1986. "The Significance of Choice." *The Tanner Lectures on Human Values* 7: 149–216.

Scanlon, Thomas. 2000. *What We Owe to Each Other*. Nachdr. Cambridge, Mass.: Belknap Press of Harvard Univ. Press.

Scanlon, Thomas. 2008. *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge, Mass: Belknap Press of Harvard University Press.

Schroeter, Laura, and Francois Schroeter. 2018. "Metasemantics and Metaethics." In *The Routledge Handbook of Metaethics*, edited by Tristram Colin McPherson and David Plunkett. Routledge Handbooks in Philosophy. New York: Routledge, Taylor & Francis Group.

Scommegna, Paola. 2013. "Exploring the Paradox of U.S. Hispanics' Longer Life Expectancy." *Population Reference Bureau* (blog). 2013. https://www.prb.org/us-hispanics-life-expectancy/.

Shelby, Tommie. 2003. "Ideology, Racism, and Critical Social Theory." *The Philosophical Forum* 34 (2): 153–88.

Shelby, Tommie. 2014. "Racism, Moralism, and Social Criticism." *Du Bois Review: Social Science Research on Race* 11 (01): 57–74. https://doi.org/10.1017/S1742058X14000010.\

Sher, George. 2006. *In Praise of Blame*. New York: Oxford University Press.

Shoemaker, David. 2011. "Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility." *Ethics* 121 (3): 602–32. https://doi.org/10.1086/659003.

Simon, Stacy. 2019. "Facts & Figures 2019: US Cancer Death Rate Has Dropped 27% in 25 Years." American Cancer Society.

Smith, Angela M. 2007. "On Being Responsible and Holding Responsible." *The Journal of Ethics* 11 (4): 465–84. https://doi.org/10.1007/s10892-005-7989-5.

Smith, Angela. 2013. "Moral Protest and Moral Blame." In *Blame: Its Nature and Norms*, edited by D. Justin Coates and Neal A. Tognazzini, 27–48. Oxford: Oxford University Press.

Spencer, Quayshawn. 2018. "Racial Realism II: Are Folk Races Real?" *Philosophy Compass* 13: 1-16. https://doi.org/10.1111/phc3.12467

Sripada, Chandra Sekhar, and Stephen Stitch. 2007. "A Framework for the Psychology of Norms." In *The Innate Mind: Volume 2: Culture and Cognition*, edited by Peter Carruthers, Stephen Laurence, and Stephen Stich, 280–301. New York, NY: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195310139.001.0001.

Sripada, Chandra. 2016. "Self-Expression: A Deep Self Theory of Moral Responsibility." *Philosophical Studies* 173 (5): 1203–32. https://doi.org/10.1007/s11098-015-0527-9.

Strawson, P.F. 1962. "Freedom and Resentment." In *Perspectives on Moral Responsibility*, edited by John Martin Fischer and Mark Ravizza, 45–68. Ithaca, NY: Cornell University Press.

Stroud, Sarah. 2006. "Epistemic Partiality in Friendship." *Ethics* 116 (3): 498–524. https://doi.org/10.1086/500337.

Sumner. 1864. "To Secure Equality Before the Law in the Courts of the United States." Senate Bill Report Senate Report No. 25. U.S. Senate.

Taylor, Paul C. 2013. *Race: A Philosophical Introduction*. Second edition. Cambridge, UK: Polity.

Tinbergen, N. 1963. "On aims and methods of ethology." Zeitschrift für Tierpsychologie 20: 410-433.

Touré. 2012. *Who's Afraid of Post-Blackness? What It's Like to Be Black Now*. New York: Atria Books.

Vargas, Manuel. 2013. *Building Better Beings: A Theory of Moral Responsibility*. 1st ed. Oxford: Oxford University Press.

Washington, Natalia, and Daniel Kelly. 2016. "Who's Responsible for This? Moral Responsibility, Externalism, and Knowledge about Implicit Bias." In *Implicit Bias and Philosophy, Volume 2: Moral Responsibility, Structural Injustice, and Ethics*, edited by Michael Brownstein and Jennifer Saul, 11–36. New York, NY: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198766179.001.0001.

Wieviorska, Michel. 1994. "Racism and Modernity in Present-Day Europe." *Thesis Eleven* 35: 51–61.

Williams, Bernard. 1985. *Ethics and the Limits of Philosophy*. Cambridge, Mass: Harvard University Press.

Williams, David R., and Selina A. Mohammed. 2009. "Discrimination and Racial Disparities in Health: Evidence and Needed Research." *Journal of Behavioral Medicine* 32 (1): 20–47. https://doi.org/10.1007/s10865-008-9185-0.

Winner, Langdon. 1980. "Do Artifacts Have Politics?" *Daedalus* 109 (1): 121–36.

Witt, Charlotte. 2011. *The Metaphysics of Gender*. Studies in Feminist Philosophy. Oxford ; New York: Oxford University Press.

Wolf, Susan. 1990. *Freedom within Reason*. New York: Oxford Univ. Press.

Womack, Lindsay, Lauren Rossen, and Joyce Martin. 2018. "Singleton Low Birthweight Rates, by Race and Hispanic Origin: United States: 2006-2016." NCHS Data Brief 306. Centers for Disease Control and Prevention.

Wylie, Alison. 2004. "Why Standpoint Matters." In *The Feminist Standpoint Theory Reader: Intellectual and Political Controversies*, edited by Sandra Harding, 339–51. New York, NY: Routledge.

Yancy, George. 2008. *Black Bodies, White Gazes: The Continuing Significance of Race*. Lanham, Maryland: Rowman and Littlefield.

Zheng, Robin. 2016. "Attributability, Accountability, and Implicit Bias." In *Implicit Bias and Philosophy, Volume 2: Moral Responsibility, Structural Injustice, and Ethics*, edited by Michael Brownstein and Jennifer Saul, 62–89. New York, NY: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198766179.001.0001.

Zheng, Robin. 2018. "What Is My Role in Changing the System? A New Model of Responsibility for Structural Injustice." *Ethical Theory and Moral Practice* 21 (4): 869–85. https://doi.org/10.1007/s10677-018-9892-8.